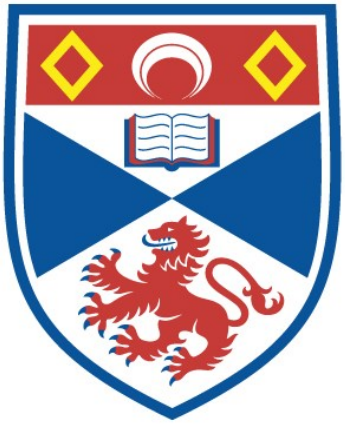
STATISTICAL ASPECTS OF THE POPULATION REGULATION OF MIGRATING BROWN TROUT, 'SALMO TRUTTA' IN A LAKE DISTRICT STREAM

Robert John Fryer

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1990

STATISTICAL ASPECTS OF THE POPULATION REGULATION OF MIGRATORY BROWN TROUT, *Salmo trutta*, IN A LAKE DISTRICT STREAM

by

Robert John Fryer

being a Thesis submitted to the University of St Andrews in candidature for the degree of Doctor of Philosophy

May 1989

Department of Agriculture and Fisheries for Scotland
Marine Laboratory
PO Box 101
Victoria Road
Aberdeen AB9 8DB

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

I declare that the work reported in this Thesis is my own and has not been accepted in any previous application for a higher degree.

Signature

SUPERVISOR'S DECLARATION

I declare that the conditions of the resolution and regulations for the degree of PhD have been fulfilled.

Signature

I was admitted as a research student at St Andrews University on 1 October 1985 and as a candidate for the PhD on 1 October 1986.

Signature

Access to the Thesis is open, subject to any regulations of the Library committee.

ABSTRACT

STATISTICAL ASPECTS OF THE POPULATION REGULATION OF MIGRATORY BROWN TROUT, *Salmo trutta*, IN A LAKE DISTRICT STREAM

by Robert John Fryer

Statistical aspects of the population regulation of a migratory brown trout population are investigated. The life cycle of the trout, the study area and the sampling routine are described in Chapter 1. Models of numerical changes in fish populations are reviewed in Chapter 2. Measures that assess the nonlinear behaviour of nonlinear regression models are described in Chapter 3. The additive error Ricker model describes the relationship between the number of 0+ parr in May/June and the number of eggs. The nonlinear behaviour of the model is investigated in Chapter 4. The parameter effects nonlinearity of the model is reduced by a reparameterisation. Chapter 5 investigates the effect of errors in the egg variable on the distributions of the least squares estimators of the additive error and the multiplicative error Ricker models. The errors-in-variables considerably increase the variances of the least squares estimators. Models of the relationships between the numbers of 0+ parr in August/September, the number of 1+ parr, the egg production of a year class and the number of eggs are developed in Chapter 6. These models account for the effect of summer drought on survival. Survival is density dependent during the first summer of the life cycle and density independent thereafter. Standard measures of nonlinearity can seriously underestimate the nonlinear behaviour of piecewise linear change-point models. New measures of nonlinearity appropriate for piecewise linear change-point models are developed in Chapter 7. Chapter 8 develops a model of the growth of brown trout fed on maximum rations as a function of time, body weight and water temperature. Chapter 9 develops a model that relates the survival rate of 0+ parr between May/June and August/September to the length distribution of the trout in May/June. The results of the Thesis are discussed in Chapter 10.

ACKNOWLEDGEMENTS

I would like to express my thanks to Richard Cormack for his invaluable guidance and advice during the course of this Thesis. I thank Malcolm Elliott for permission to use the Black Brows Beck data and gratefully acknowledge the privilege of being associated with such an important and elegant ecological study. I would also like to thank Margaret Hurley for her help and encouragement over the past three years and in particular for the great time I had whilst at the Freshwater Biological Association. I am also grateful to Tony Irish for help with various computing problems and to Lesley Davidson and Patricia Bates for help with word processing this Thesis.

This work was financed by a NERC research studentship.

CONTENTS

APPENDIX 5

CHAPTER 1 INTRODUCTION

The most detailed investigation of long term population changes in brown trout, *Salmo trutta* L., is Elliott's study of the migratory brown trout population in Black Brows Beck, a small stony stream in the English Lake District (Elliott, 1984a,b, 1985a,b,c,d, 1986, 1987, 1988). To date, Elliott has examined numerical changes and population regulation at different stages of the life cycle (Elliott, 1984a, 1985a,b,d), variations in growth, production and biomass at different stages of the life cycle (Elliott, 1984b, 1985c,d) and the spatial distribution and behavioural movements of the population (Elliott, 1986). Population regulation and growth in Black Brows Beck have also been compared with population regulation and growth of a neighbouring brown trout population in Wilfin Beck (Elliott, 1987, 1988).

One of the features of Elliott's study is his sampling routine to estimate population numbers (described in Section 1.2). The life cycle of the trout was divided into short stages and population estimates were obtained at each stage over a long period of time (1967-1988). Many of these estimates were obtained with a very small measurement error. By investigating survival rates between one stage and the next, it is possible to identify important factors affecting survival and to develop quite complex models of the dynamics of the trout population. This Thesis investigates some statistical aspects of the study of population regulation in Black Brows Beck. In particular, some properties of the existing models of population regulation are examined and new models of population regulation are developed. First, this Chapter describes the life cycle of the trout, the study section of the stream and the sampling

routine, summarises Elliott's results on population regulation and outlines the work in this Thesis.

1.1 LIFE CYCLE (ELLIOTT, 1984a,b, 1985a,d)

Spawning occurred in November and early December (see Fig. 1). The eggs were laid in redds, depressions cut in the gravel bed of the stream by the spawning female, and were then covered with more gravel. No female was observed to cut more than one redd.

The fish hatched around the following February. They remained in the gravel nest feeding off the yolk-sac until April or May, when they emerged from the gravel and started to feed. Most fish remained in fresh water for two years and then migrated to sea. However, a few fish remained an extra year in fresh water and migrated at the start of their fourth year.

A large number of mature males returned to spawn in November or December after spending only one summer in the sea. A few of the females who remained in fresh water for three years also returned to spawn after just one summer at sea. Other trout returned after spending two summers in the sea. Having spawned, the trout migrated downstream; a few trout returned to spawn again the following year.

Migratory brown trout are known as

i) alevins in the period between hatching and emerging from the gravel nest,

ii) fry in the short transition period between emerging from the gravel nest and the time when the yolk-sac has been fully absorbed,

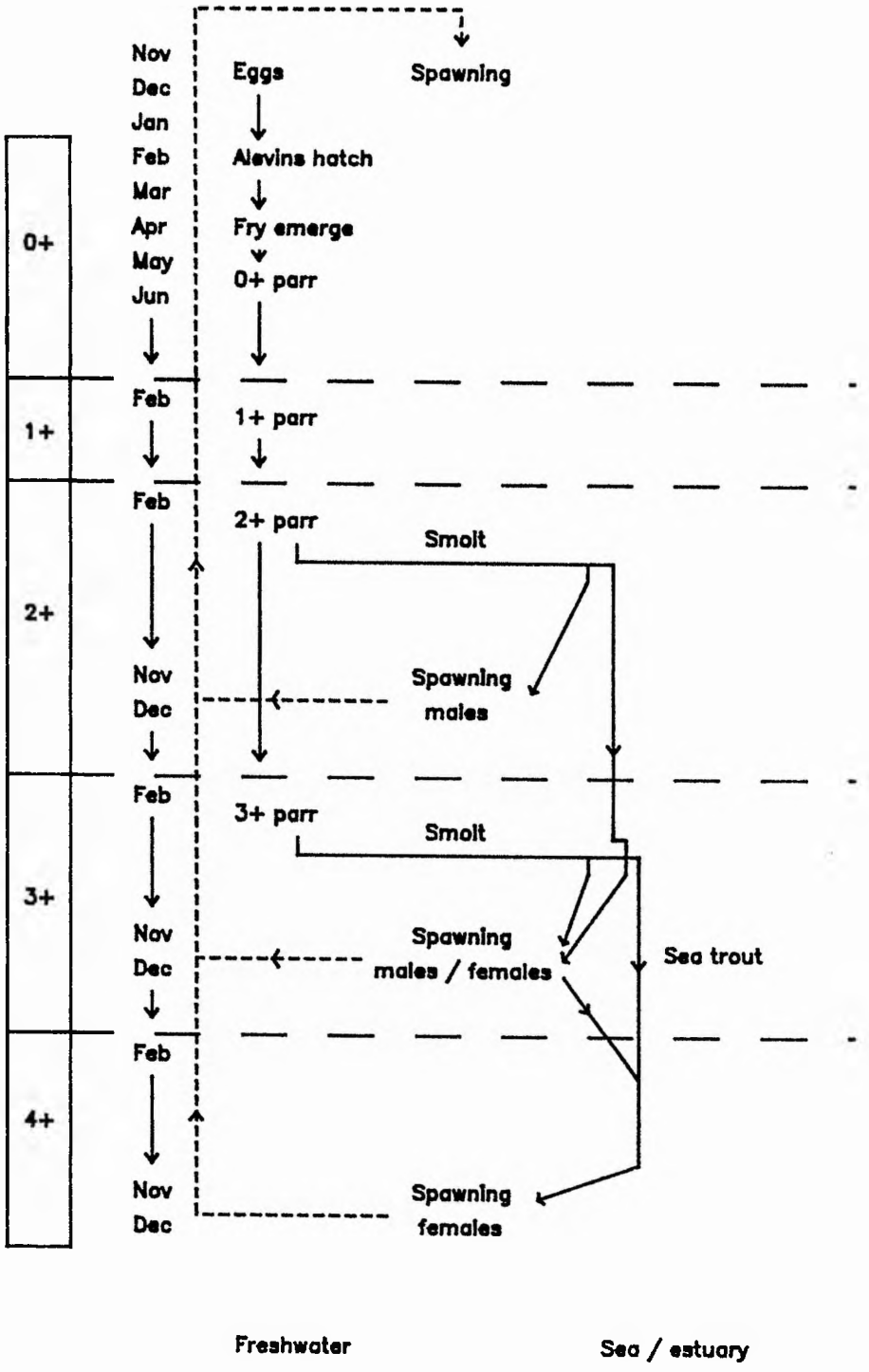


Fig. 1 The life cycle of the Black Brows Beck trout population.

iii) parr in the period from full absorption of the yolk-sac to the time when the trout prepares to migrate,

iv) smolts in the period when seaward migration occurs.

The ages of the trout are described as follows: 0+ trout are less than one year old, 1+ trout are between one and two years old, etc. The age designation changes in February, since this was the month in which most eggs hatched. Finally, a year class is known by the year in which the eggs hatched; for example, the eggs of the 1967 year class were laid in November/December 1966 and hatched around February 1967.

1.2 SAMPLING ROUTINE (ELLIOTT, 1984a,b, 1985a,d)

The study section of Black Brows Beck was 120 m long, with a mean width of 0.8 m and a water depth that generally varied between about 6 cm in riffles to about 40 cm in deep pools. (A more detailed description of the location, vegetation, etc of the stream is given in Elliott, 1984a). The study section was divided into two parts. The upper section (of length 75 m and area 60 m^2) was used for all estimates of population numbers. Throughout this Thesis, the study of population regulation is implicitly the study of changes in the numbers of trout in the upper study section of Black Brows Beck. The lower study section was used to investigate other aspects of the population, such as the number of eggs per redd.

The sampling routine each year (November 1966-1988) was as follows. The redds cut in the (upper) study section were counted in November/December. The number of eggs laid was then estimated by multiplying the number of redds by estimates

of the average numbers of eggs per redd (a full description is given in Section 5.2). The number of females that had returned to spawn was also estimated from the number of redds.

The alevins were sampled approximately 10 days before their emergence from the gravel, at time t_0 say. As the emergence date depended on the water temperature throughout the winter, t_0 varied between years, but was usually sometime in April. The study section was divided into 10 subsections, each of area 6 m^2, and three random samples, each covering an area 0.09 m^2, were taken in each subsection. The number of alevins in the upper study section was estimated by dividing the number of alevins sampled by the proportion of the area of the stream that was sampled.

The 0+ parr were sampled at time t_1 (approximately 32 days after t_0) in May/June and at time t_2 (approximately 125 days after t_0) in August/September. The study section was divided into six subsections, each of area 10 m^2. Block nets were placed at the top and bottom of each subsection (so that trout could neither enter nor leave) and each subsection was then repeatedly electro-fished until no fish were caught in two successive fishings. The numbers of 1+ parr and 2+ parr were obtained at the same time (although there were never many 2+ parr, since most trout had migrated by this stage). Since the entire upper study section was electro-fished so thoroughly, the numbers of parr were estimated very accurately (perhaps - one or two fish).

Let t_3 and t_4 be the dates of the May/June and August/September samplings of the 1+ parr. For each year class i, let

i) E_i be the number of eggs laid,

ii) R_{0i} be the number of alevins at time t_0,

iii) R_{1i}, R_{2i} be the numbers of 0+ parr at times t_1, t_2 respectively,

iv) R_{3i}, R_{4i} be the numbers of 1+ parr at times t_3, t_4 respectively,

v) E^*_i be the egg production of the ith year class (ie the number of eggs laid by trout that hatched in year i).

This Thesis investigates population regulation in the year classes 1967-1984. Estimates of E, $R_0,...,R_4$, E^*, the numbers of 2+ parr in May/June and August/September and the number of females that returned to spawn, for each of these year classes are given in Table 1.

1.3 SURVIVAL (ELLIOTT, 1984a, 1985a,b,d)

This Section summarises Elliott's results on population regulation in Black Brows Beck.

The estimates of the number of eggs E and the number of alevins R_0 are very similar for each year class (Table 1). Hence, there is negligible mortality between the time the eggs are laid and t_0.

The relationships between $R_1,...,R_4$, E^* and E (or equivalently, between $R_1,...,R_4$, E^* and R_0) are more complicated. In a comparison of six stock-recruitment models, the best general model for describing these relationships is found to be the additive error Ricker model

$$R_i = a_1 E_i \exp(-b_1 E_i) + \epsilon_i, \qquad i = 67...84, \tag{1.1}$$

TABLE 1

Estimates of population numbers

Year class	E	R_0	R_1	R_2	R_3	R_4	2+ parr		Spawning females	E*
							M/J	A/S		
1967	4652	4622	346	116	35	28	4	2	6	4644
1968	2272	2267	409	92	41	14	5	5	4	2888
1969	1756	1622	403	74	27	30	4	0	5	3610
1970	2584	2311	460	115	51	47	8	0	7	7958
1971	4644	4333	308	110	46	44	7	0	7	7334
1972	2888	2489	439	112	48	45	10	0	6	5784
1973	3610	3334	412	119	45	44	8	0	5	4750
1974	7958	8400	132	74	30	25	1	0	2	2068
1975	7334	7444	182	88	42	23	5	2	1	1034
1976	4234	4489	354	91	37	38	2	4	3	3822
1977	6300	6222	241	99	29	31	1	0	5	4546
1978	2068	2000	429	105	46	38	2	0	7	7958
1979	1034	1089	334	73	30	25	3	0	3	2478
1980	722	844	286	61	27	21	4	0	3	2166
1981	7646	7933	178	87	44	37	2	0	3	4134
1982	7958	7978	158	81	41	14	0	0	0.33	517
1983	2478	2400	425	43	12	2	0	0	1	1034
1984	2166	2200	394	15	13	12	0	0	4	5684

where

i) R represents any one of the life stages $R_1,...,R_4$ (or E^*),

ii) a_1, b_1 are parameters of the model (whose values depend on the life stage R),

iii) the suffix i runs through the year classes 1967-1984,

iv) the additive random errors ε_i, $i = 67...84$, are independent and normally distributed with zero mean and constant variance σ^2 (ie $\{\varepsilon_i\}$ are NID$(0, \sigma^2)$).

Model (1.1) is an excellent fit to the R_1 data. However, the model fits the later life stages less well. This is because periods of summer drought greatly reduced survival between t_1 and t_2 and between t_3 and t_4 in some year classes, and the Ricker model fails to account for this drought effect. The multiplicative error Ricker model

$$R_i = a_1 E_i \exp(-b_1 E_i) \exp(\epsilon_i), \qquad i = 67...84, \tag{1.2}$$

where $\{\varepsilon_i\}$ are NID$(0, \sigma^2)$, gives similar fits to the data as the additive error Ricker model. (The Ricker models are described in more detail in Chapters 2 and 4.)

Drought was the only environmental factor found to affect survival. In particular, high spates, the density of older trout in the stream and the density of other fish species (bullheads, *Cottus gobio* L. and eels, *Anguilla anguilla* (L.)) in the stream had no obvious effect on survival.

Survival is density dependent between t_0 and t_1 and between t_1 and t_2. After t_2, survival is density independent; the numbers of trout at times after t_2 are proportional (neglecting the effect of summer droughts) to R_2 (ie R_3 and R_4 are proportional to R_2).

It is important to note that migration in and out of the study section could occur. Losses between t_0 and t_2 and between t_3 and t_4 were due almost entirely to mortality; however, losses between t_2 and t_3 were due to both mortality and some downstream migration. The number of trout in the study section could also increase with time through immigration (*cf* R_3 and R_4 in the 1969 year class). Throughout this Thesis, terms such as losses, survival, etc implicitly refer to the net effect of mortality and migration.

1.4 THE SCOPE OF THIS THESIS

A wide variety of models have been developed to describe numerical changes in fish populations. A review of these, including a more detailed description of the Ricker models, is given in Chapter 2.

The additive error Ricker model (1.1) is fitted by unweighted nonlinear least squares regression. Standard inference is then based on the assumption that the "nonlinear behaviour" of the model is small; ie that the model behaves very like a linear model. A review of measures that assess the nonlinear behaviour of a nonlinear regression model is given in Chapter 3. In Chapter 4, the nonlinearity of the Ricker model is investigated, with particular reference to the R_1 data set; the model is reparameterised to reduce the nonlinear behaviour.

Both Ricker models (1.1) and (1.2) are fitted by least squares regression. The additive error model is fitted by nonlinear least squares. The multiplicative model is fitted by linearising the model to

$$\log(R_i) = \log(E_i) + a_2 - b_1 E_i + \epsilon_i, \qquad i = 67...84, \tag{1.3}$$

where $a_2 = \log(a_1)$ and $\{\epsilon_i\}$ are NID(0, σ^2), and then using linear least squares. Both methods, and inference based on them, assume that the independent variate E is known without error (or at least that errors in the independent variate are negligible). This assumption is not strictly valid for Black Brows Beck. Chapter 5 investigates the effect of the errors in the egg variable on the distribution of the least squares estimator of each model.

A measure of summer drought based on the water level of Esthwaite Water, a lake close to Black Brows Beck, is developed in Chapter 6. This measure is incorporated into the additive error Ricker model to produce models that give good descriptions of the relationships between $R_2,...,R_4$, E^* and E, even in year classes affected by summer drought.

Density dependent survival occurs between t_0 and t_1 and between t_1 and t_2. The density dependence is such that the relationships between R_1 and E and between R_2 and E are well described by comparatively simple models (Chapters 4, 6). However, the relationship between R_2 and R_1 is more complicated; specifically, the survival rate between t_1 and t_2 depends on more than just the number of trout at t_1. Chapters 7, 8 and 9 lead to the development of a model of survival between t_1 and t_2. In Chapter 7, the nonlinearity of piecewise linear models with unknown

change points is discussed. The standard analytical measures of nonlinearity described in Chapter 3 can not be applied to these models because the derivatives of the models with respect to their parameters are not continuous. A new measure of nonlinearity that can be applied to these models is developed. In Chapter 8, the new nonlinearity measures are used in the analysis of the results of a series of experiments on the growth of brown trout fed on maximum rations (Elliott, 1975); a model describing the growth of trout as a function of time, body weight and water temperature is developed. Finally, in Chapter 9, the growth model is used in the development of a model that relates the survival rate between t_1 and t_2 to the length distribution of the trout at t_1.

The results of this Thesis are discussed in Chapter 10.

CHAPTER 2 MODELS OF FISH POPULATION DYNAMICS

2.1 INTRODUCTION

Since the early works of Ricker (1954, 1958), Beverton and Holt (1957) and Schaefer (1954, 1957), there has been a steady development of models that describe the dynamics of fish populations. This has been particularly motivated by the collapse of many of the world's exploited fish stocks, such as the North Sea herring; it is important to be able to model populations in such a way that the effect of various fishing policies can be accurately predicted, so that fish stocks and markets can be preserved.

This Chapter reviews models of fish population dynamics. Section 2.2 describes simple production models, which give very rough descriptions of the dynamics of a population. Stock-recruitment models, which relate recruitment to a population to the spawning stock size, are described in Section 2.3. More complicated models, which involve the age-structure of a population, are described in Sections 2.4 and 2.5. Finally, in Section 2.6, statistical problems that arise in modelling fish population dynamics are discussed.

2.2 PRODUCTION MODELS

The most basic model of fish population dynamics is the simple production or surplus-yield model, in which the growth of a population at time t is determined by the size of the population at that time. This is generally expressed as

$$\frac{dP(t)}{dt} = f(P(t)) - qe(t)P(t) \qquad (2.1)$$

where

i) $P(t)$ is the size of the population at time t (in numbers of fish or biomass of the population),

ii) $f(.)$ is the "production" function,

iii) $e(t)$ is the fishing effort at time t,

iv) q is the catchability coefficient, which measures the "efficiency" of the fishing effort (ie as q increases, more fish are caught per unit effort).

In (2.1), the term $-qe(t)P(t)$ represents the effect of mortality due to exploitation. The production function $f(.)$ represents the net effect of recruitment, natural mortality and, (if P is expressed in terms of the stock biomass,) the increase in weight of fish within the population. However, in general, $f(.)$ does not contain terms that explicitly account for any of these effects.

Production functions generally assume that growth is approximately proportional to P when P is small and that growth is restricted when P is large. Some common examples are

i) $f(P) = aP - bP^2$, (2.2a)

logistic production function (Schaefer, 1954),

ii) $f(P) = aP - bP^m$, (2.2b)

(Pella and Tomlinson, 1969),

iii) $f(P) = aP - bP\log(P)$, (2.2c)

(Fox, 1970),

where $a > 0$, $b > 0$ and $m > 1$ are parameters of f; examples of i) and iii) are shown in Figure 2.1.

Production models are very simple models and are generally regarded as deterministic. They are useful, because they can be applied to populations for which very little data are available. In particular, parameter estimates can often be obtained using only catch and effort data. For example, model (2.1) with the logistic production function (2.2a) (known as the Schaefer model) can be rearranged (with suitable approximations) to give

$$\log\left(\frac{\overline{U}_{i+1} + \overline{U}_i}{\overline{U}_i + \overline{U}_{i-1}}\right) = a - q\bar{e}_i - \frac{b}{q}\overline{U}_i, \quad i = 1...n \tag{2.3}$$

(Schaefer, 1957),

or

$$\log\left(\frac{\overline{U}_{i+1}}{\overline{U}_i}\right) = a - q\left(\frac{\bar{e}_i + \bar{e}_{i+1}}{2}\right) - \frac{b}{q}\left(\frac{\overline{U}_i + \overline{U}_{i+1}}{2}\right), \quad i = 1...n \tag{2.4}$$

(Schnute, 1977),

where

i) $\bar{e}_i$ is the average fishing effort in year i,

ii) $\overline{U}_i$ is the average catch per unit effort in year i.

Estimates of the parameters a, b and q are then obtained by ordinary least squares regression. (This estimation procedure implies there is a stochastic component in

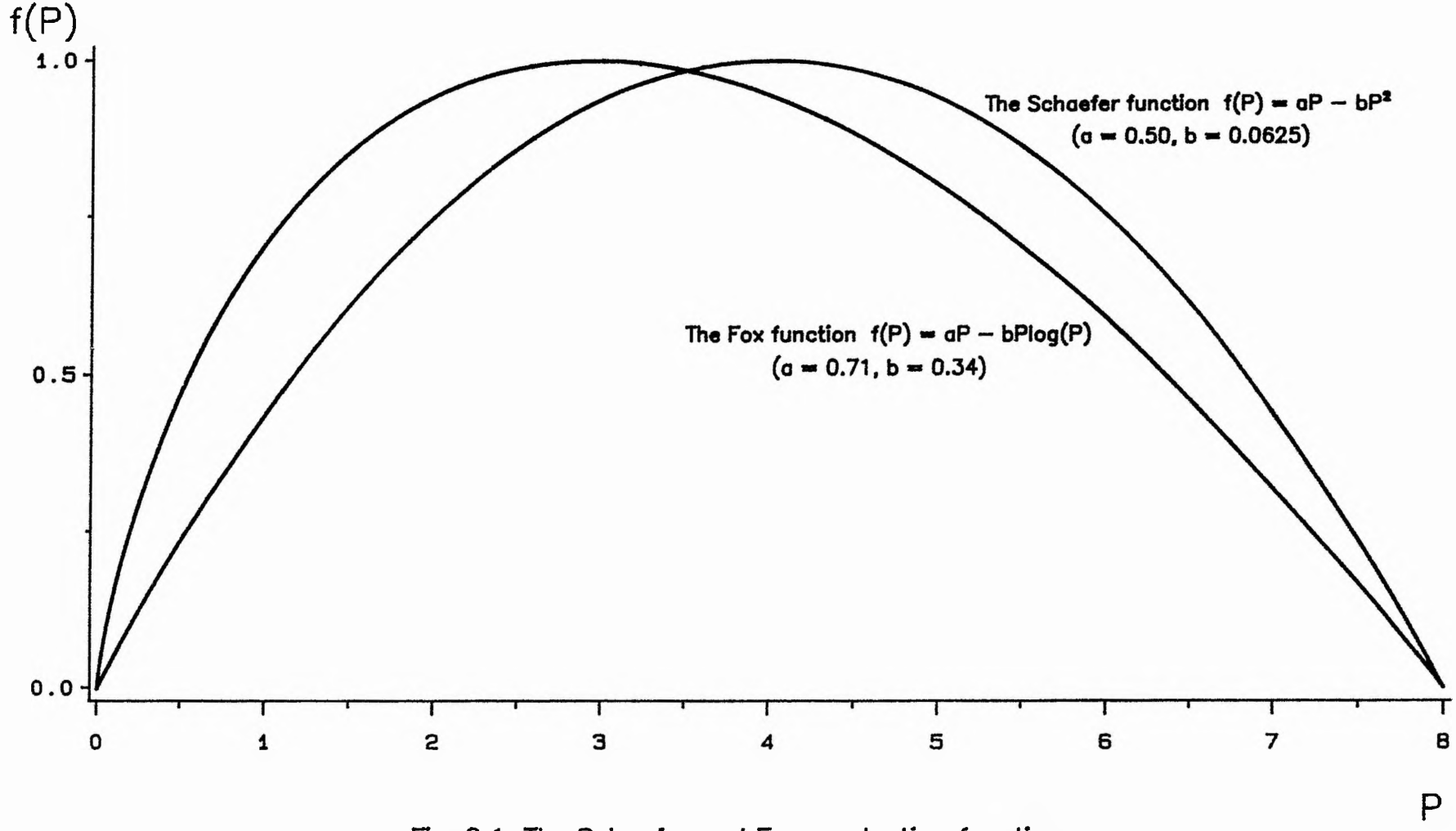


Fig. 2.1 The Schaefer and Fox production functions
(standardised to have the same maximum values).

the model; however, once the parameter estimates have been obtained, random variation is generally ignored. Statistical problems of estimation that arise through using formulations such as (2.3) and (2.4) are discussed in Section 2.6.)

Model (2.1) assumes that recruitment to the population is instantaneous. This assumption can be relaxed to allow for delayed recruitment, by adding a time lag to the production function. For example, the Rhode Island inshore lobster fishery is well described by the delayed recruitment model

$$\frac{dP(t)}{dt} = \alpha P(t) - bP(t)^2 - cP(t-5) - qe(t)P(t)$$

where the time lag of five years corresponds to the average age of recruitment to the fishery (Marchesseault, Saila and Palm, 1976). However, explicit terms for natural mortality, recruitment etc must be introduced into the production function to generalise model (2.1) further.

2.3 STOCK-RECRUITMENT MODELS

The period between spawning and recruitment is generally regarded as the stage in the life cycle in which most random and density dependent effects occur. Consequently, much emphasis has been placed on the development of stock-recruitment models, in which the recruits R (in numbers or biomass) are related to the spawning stock S (in numbers of adults, biomass or numbers of eggs), by

$$R = f(S) \qquad (2.5)$$

in conjunction with some form of random variation; $f(.)$ is known as the stock-recruitment function. Some examples are

i)	$R = aS\exp(-bS)$,	(Ricker, 1954),	(2.6a)
ii)	$R = aS/(1 + bS)$,	(Beverton and Holt, 1957),	(2.6b)
iii)	$R = aS/(1 + (bS)^c)$,	(Shepherd, 1982),	(2.6c)
iv)	$R = aS(1 - bcS)^{1/c}$,	(eg Schnute, 1985),	(2.6d)

Stock-recruitment models have also been used in many other biological contexts (see reviews by Elliott, 1985b; May and Oster, 1976).

The stock-recruitment functions (2.6) all describe particular forms of density dependent recruitment. For example, consider the Ricker and Beverton and Holt models (2.6a,b). At low stock densities, R varies almost linearly with S; however, as S increases, density has an increasingly inhibitory effect on recruitment. In the Beverton and Holt model, as S increases, R tends to the limiting value of a/b. In the Ricker model, the density dependent effect is greater; as S increases, R increases to some maximum value and then decreases (Fig. 2.2). Three-parameter stock-recruitment functions, such as (2.6c) and (2.6d), can typically describe both asymptotic and dome-shaped relationships.

The parameters of stock-recruitment functions can often be interpreted biologically. For example, in all four functions (2.6), the parameter a is a measure of density independent recruitment; that is, at low stock sizes (when the density dependent effects are negligible), a gives the approximate increase in R for each unit increment in S. In the Ricker and Beverton and Holt models (2.6a,b), b is a measure of density

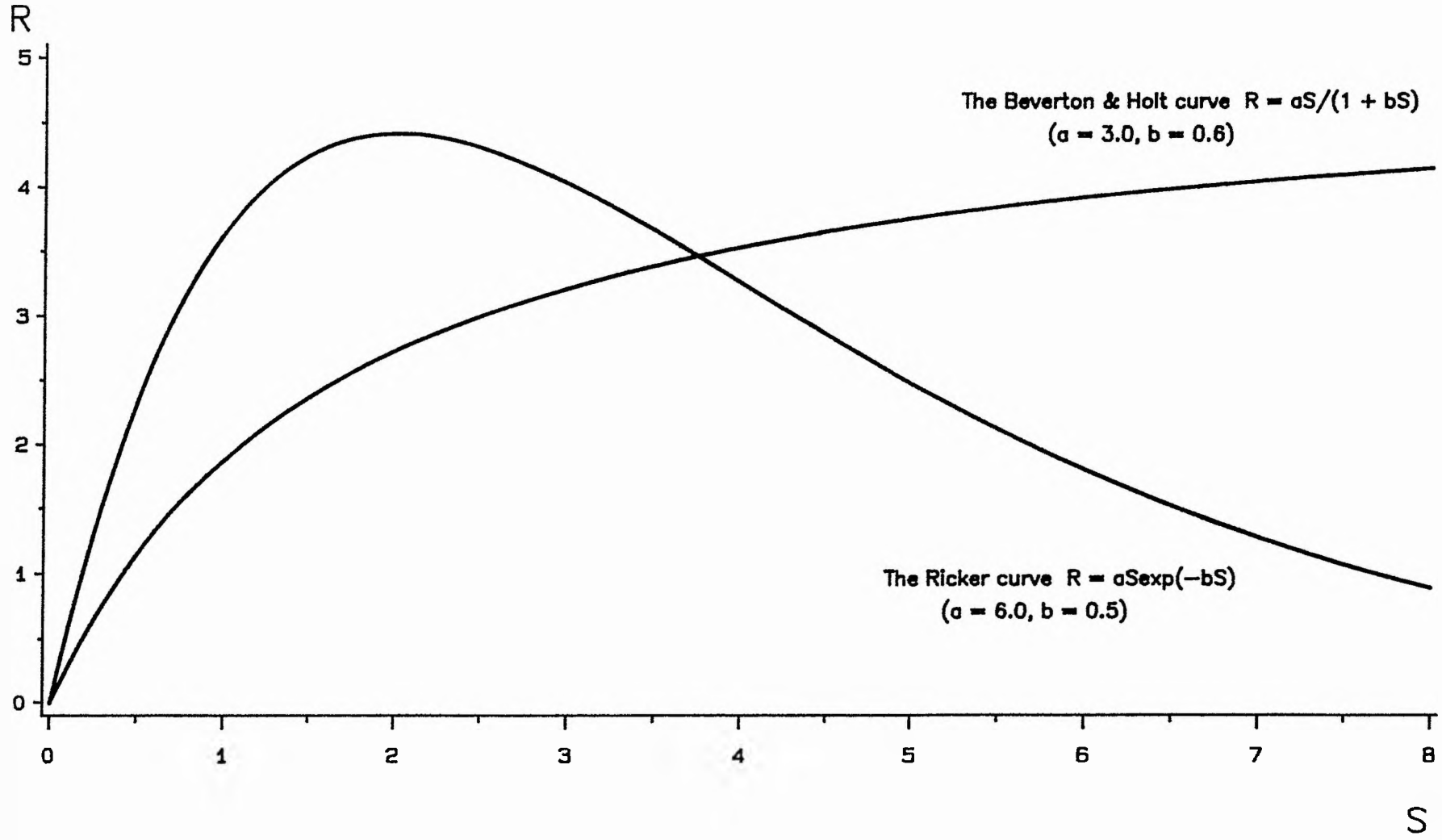


Fig. 2.2 The Ricker and Beverton & Holt stock–recruitment curves.

dependent recruitment; as b increases, density dependent effects become more important at lower stock densities (for example, see Fig. 2.3). In the Ricker model, recruitment is maximised when $S = 1/b$; thus stock sizes over $1/b$ have a serious inhibitory effect on survival. Similarly, in the Shepherd model (2.6c), $1/b$ is the threshold stock size above which density dependent factors dominate; c gives the degree of density dependence.

The Ricker and Beverton and Holt models can both be derived theoretically by making specific assumptions about the instantaneous natural mortality rates of young fish. However, their use is generally based on purely empirical grounds. Most three-parameter stock-recruitment functions were developed to represent a wide range of spawner-recruit behaviour; an exception is the stock-recruitment function

$$R = \left(\frac{\alpha \exp(bS)}{S} + c\right)^{-1}$$

which was developed on a bioenergetics principle (Ware, 1980).

The choice of an appropriate random component to use with (2.5) depends on the population in question. Two forms which are often used are

i) an additive normal error with constant variance

$$R_i = f(S_i) + \epsilon_i, \qquad i = 1 \ldots n,$$

where $\{\epsilon_i\}$ are NID$(0, \sigma^2)$,

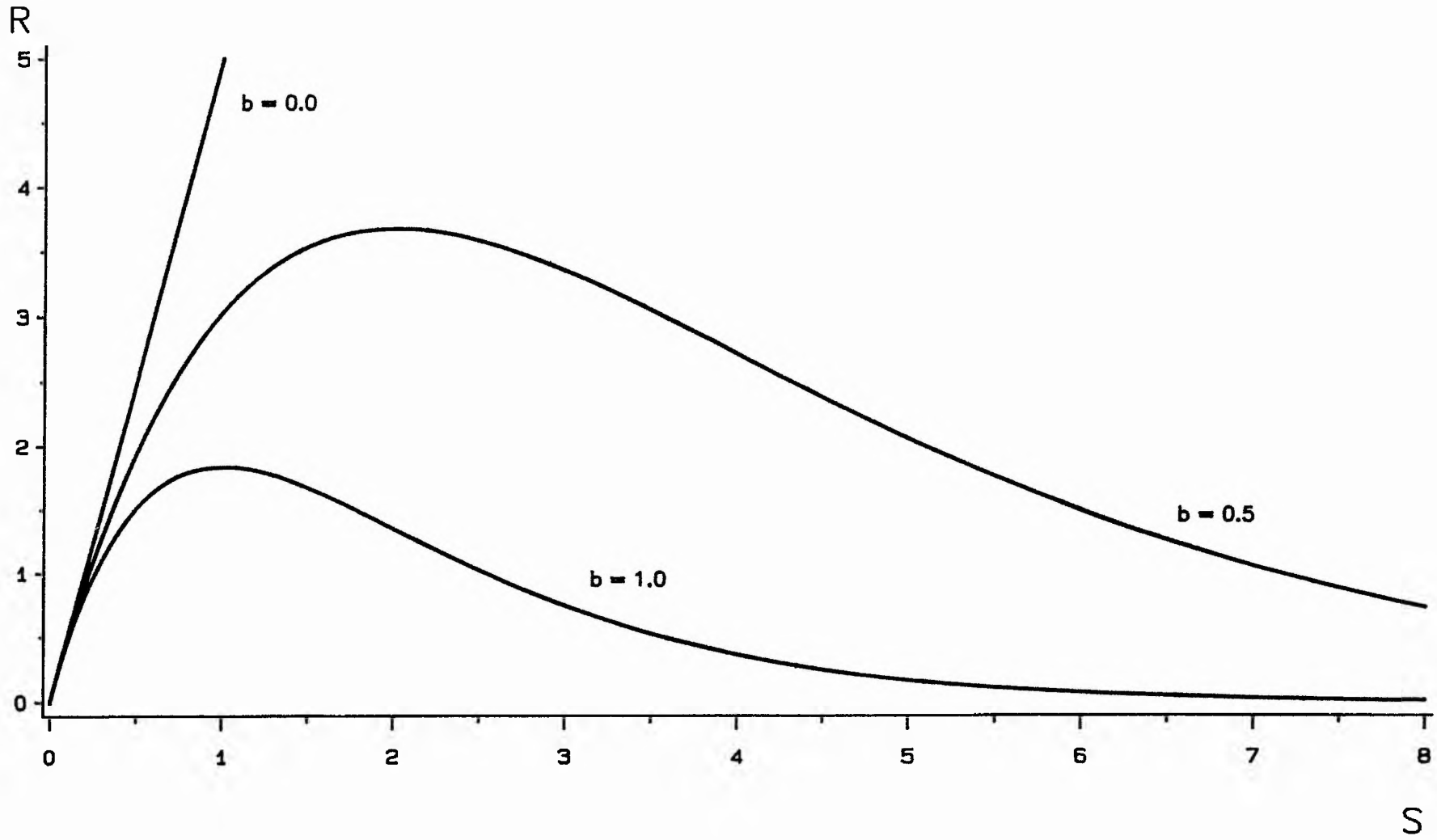


Fig. 2.3 Ricker curves for different values of b (a = 5.0).

ii) a multiplicative normal error with constant variance

$$R_i = f(S_i)\exp(\epsilon_i), \quad i = 1 \ldots n,$$

where $\{\epsilon_i\}$ are NID(0, σ^2).

Estimates of $\{R_i\}$ and $\{S_i\}$ are generally required to estimate the parameters of stock-recruitment models (although when a stock-recruitment model forms part of a more complicated model (see, for example, Section 2.4), the parameters can sometimes be estimated using only catch and effort data). The method of parameter estimation depends on the form of the random component.

It is very rare that the dynamics of a population can be described by a stock-recruitment model alone, since it is usually also necessary to incorporate the effect of exploitation and, in many populations, age-structure. However, stock-recruitment models describe a very important stage in the life cycle and form an integral part of most of the more complicated models described in Sections 2.4 and 2.5.

2.4 DELAY-DIFFERENCE POPULATION MODELS

2.4.1 The Deriso Model

A serious limitation of the production models in Section 2.2 is that they ignore the age-structure of a fish population. An important model that incorporates age-structure, whilst still retaining only a relatively small number of parameters, is the delay-difference population model of Deriso (1980).

Deriso considers an exploited, seasonally breeding population in which

i) recruits join the catchable population at the start of the year, k years after being spawned,

ii) fishing then occurs throughout the year,

iii) spawning occurs at the end of each year; the stock that remains (uncaught) at the end of the year "generates" the recruits that join the catchable population k years later.

The population dynamics are described in terms of the biomasses of the catchable adults (B_i) at the start of each year, the catches (in weight) (C_i) during each year and the spawning stock biomasses ($S_i = B_i - C_i$) at the end of each year. If N_{ai} is the number of fish of age a ($a \geq k$) at the start of year i, and w_{ai} is the average weight of fish of age a ($a \geq k$) during year i, then approximately,

$$B_i \cong \sum_{a \geq k} w_{ai} \; N_{ai} \tag{2.7}$$

It is assumed that the weights (w_{ai}) are related by

$$w_{ai} = w_{k,i-(a-k)} \left(\frac{1 - \rho^{1+(a-k)}}{1 - \rho} \right) \tag{2.8}$$

where ρ ($0 \leq \rho < 1$) is a growth coefficient. This, and a few other assumptions, enable the weights (w_{ai}) and the population numbers (N_{ai}) to be eliminated from (2.7). The biomass in year i+1 can then be expressed in terms of the biomasses in previous years by the Deriso delay-difference model

$$B_{t+1} = (1+\rho)lS_t - \rho l(S_t/B_t)S_{t-1} + f(S_{t+1-k}) \qquad (2.9)$$

where

i) l is the annual natural survival fraction for catchable adults; ie the proportion of adults that survive the year if there is no exploitation,

ii) $f(.)$ is a stock-recruitment function.

The parameters of (2.9) are generally estimated using catch and effort data. Confidence regions for the parameters are often very large, due to the large number of parameters; (if a flexible three-parameter stock-recruitment function is used, six parameters are estimated (l, ρ , q (the catchability coefficient) and the three recruitment parameters). However, this problem is reduced if some of the parameters, such as the growth coefficient, are estimated from data independent of the catch and effort data.

The Deriso model (2.9) has a number of limitations. First, many assumptions are still required to derive the model. For example, it is assumed that each year class is equally catchable, that natural survival is age and year independent and that fecundity is a function of the total stock biomass. Secondly, the model is often poor at predicting the effect of exploitation; simulations have shown that simple production models, based on *ad hoc* assumptions, often give as good predictions of future catches as the Deriso model, even when the Deriso model is used to generate the data (Roff, 1983; Ludwig and Walters, 1985). However, the model has a sound theoretical basis, since it is derived from an age-structured formulation and has biologically interpretable parameters.

2.4.2 Other Delay-Difference Models

The Deriso model is also used in Generalised Stock Reduction Analysis (Kimura, Balsiger and Ito, 1984). It is assumed that the stock biomasses (B_i) are related by (2.9). In addition, the catches in year i are assumed to be given by

$$C_i = B_i F_i \{1-\exp(-M-F_i)\}/(M+F_i)$$

where

i) M is the instantaneous natural mortality rate (year independent),

ii) F_i is the instantaneous fishing mortality rate in year i.

The model has more parameters than there are years of data (since each fishing mortality F_i is one parameter of the model). Hence, a set of parameter estimates can be found that exactly explain the data. This set is then reduced to just one solution using information and intuition independent of the catch data. The advantage of this technique is that it is not necessary to specify any random component within the model, since the equations can be solved exactly. This is also a great limitation, since fish populations are inevitably subject to variation and a realistic model must try to reflect this.

Schnute (1985) generalised the Deriso model to produce a large class of models with the catch equation

$$\frac{C_{i+1}}{(1-\varnothing_{i+1})} = \frac{(1+\rho)\tau_i C_i}{(1-\varnothing_i)} - \frac{\rho\tau_i\tau_{i-1}C_{i-1}}{(1-\varnothing_{i-1})} + V_{i+1-k}\, f_{i+1}\left(\frac{\varnothing_{i+1-k}C_{i+1-k}}{1-\varnothing_{i+1-k}}\right)$$

$$- \rho\tau_i v_{i-k} f_i \left(\frac{\varnothing_{i-k}C_{i-k}}{1-\varnothing_{i-k}}\right) \qquad (2.10)$$

where

i) $\emptyset_i$ (τ_i) are annual fishing (total) survival fractions in year i,

ii) ρ, V_i, ν_i are weight parameters in year i,

iii) f_i (.) is the stock-recruitment function in year i;

for a complete explanation of the parameters, see Schnute. For any particular fishery, model (2.10) is simplified by making assumptions based on the modeller's independent knowledge of the population. Taking an extreme example, if all the parameters are year independent, all fish are of a constant weight V and a constant number of recruits R join the catchable population each year, then (2.10) reduces to

$$C_{i+1} = \tau C_i + (1 - \emptyset)RV$$

Schnute suggests ways of reducing (2.10) to a manageable form and ways of introducing stochastic variation into the model. The model is the most general framework produced to date that provides a good compromise between the over simplistic production model and the full age-structured model.

2.5 THE FULL AGE-STRUCTURED MODEL

2.5.1 The General Model

The next stage in model complexity is the full age-structured model. The population is divided into age-classes, corresponding to ages k (the age of recruitment), k+1, k+2, etc. The number of fish of age a+1 in year i+1 is related to the number of fish of age a in year i by

$$N_{a+1,t+1} = N_{at}\tau_{at}(N_t), \qquad a \geq k, \tag{2.11a}$$

where

i) N_{ai} is the number of fish of age a in year i,

ii) $N_i = (N_{ki}, N_{k+1,i},...)'$ is the vector of population numbers in year i,

iii) $\tau_{ai}(.)$ is the annual survival fraction of fish aged a in year i; this is often taken to be independent of N_i and given by

$$\tau_{at} = \exp(-(M_a + q_a e_t))$$

The number of recruits in year i is given by

$$N_{kt} = f(N_{t-k}), \tag{2.11b}$$

where $f(.)$ is a stock-recruitment function; typically, f reduces to a function of either

i) the number of adults: $f(N_i) = f(\sum_{a \geq k} N_{ai})$,

ii) the stock biomass: $f(N_i) = f(\sum_{a \geq k} w_{ai} N_{ai})$,

iii) the number of eggs: $f(N_i) = f(\sum_{a \geq k} f_{ai} N_{ai})$, where f_{ai} is the fecundity of fish aged a in year i.

Stochastic variation is introduced at any stage of the model. However, it is usually assumed that the major source of variation is the stage between spawning and recruitment; generally, a stochastic component is introduced into (2.11b) and equations (2.11a) are regarded as deterministic.

Model (2.11) can represent many different types of biological behaviour. However, the model has a large number of parameters, so that a wide variety of auxiliary information is required to fit the model to real data. Thus, although age-structured models have been applied to a few populations - for example, a Striped Bass population (Cohen, Christensen and Goodyear, 1983) - their use has generally been restricted to more theoretical studies. These can be divided into three categories. First, simulations of model (2.11) are used to project populations into the future, so that the long term effect of various fishing policies can be observed; an example is the model of Pacific Salmon developed by Larkin and Hourston (1964).

Secondly, it is assumed that

i) model (2.11) is deterministic,

ii) the fishing effort e is assumed constant,

iii) the annual survival fraction is year independent and is expressed as a function of the fishing effort as

$$\tau_{at} = \tau_a(e) = \exp(-M_a - q_a e)$$

The steady-state population vector

$$N_e = (N_{ke}, N_{k+1,e}, \ldots) \qquad (2.12)$$

satisfying

$$N_{a+1,e} = N_{ae}\exp(-M_a - q_a e), \qquad a \geq k$$

$$N_{ke} = f(N_e)$$

is then found (as a function of the fishing effort) and optimal fishing strategies (eg fishing efforts that maximise the steady-state yield) are investigated; for example, see Reed (1980), Getz (1980).

The third approach is of particular interest. The steady-state population vector (2.12) is again found. However, it is now assumed that there is a stochastic component in the stock-recruitment relationship which causes the population vector to vary about the steady-state. Expressions for the variances of population numbers, biomass and yield are obtained using either a frequency domain analysis (Horwood and Shepherd, 1981; Horwood, 1982, 1983) or a time domain analysis (Reed, 1983; Getz, 1984). It has been shown that in some cases, a small reduction in fishing effort can result in a large decrease in the variance of the yield, whilst only slightly decreasing the mean yield (Horwood, 1982). Further, it has been shown that the estimators of the variances of population numbers, yield, etc can be very biased if the age-structure of a population is ignored. For example, at high fishing efforts, the variance of yield of North Sea herring can be underestimated by as much as 30% if no account is taken of the age-structure of the population (Horwood, 1982).

2.5.2 A Probability Transition Matrix Model

It is often difficult to choose an appropriate stock-recruitment model to use in the age-structured model (2.11) (or in the delay-difference models in Section 2.4) due

to the large amount of random variation in most fisheries data (see Section 2.6). This problem is avoided by the probability transition matrix approach of Getz and Swartzman (1981). They consider stock-recruitment data as points (S_{i-k}, N_{ki}) in a plane, where

i) S_{i-k} is the spawning stock in year i - k (in numbers of adults, biomass or numbers of eggs),

ii) N_{ki} is the number of recruits (ie of age k) in year i.

Upper bounds, $\hat{S}$ and $\hat{N}_k$, are placed on the spawning stock size and the number of recruits respectively. The interval [0, $\hat{S}$] is divided into m_s equal subintervals, each representing one stock class. Similarly, the interval [0, $\hat{N}_k$] is divided into m equal subintervals, each representing one recruit class. The stock-recruitment relationship is then described by a probability transition matrix T with elements

$T_{j_s j}$ = Pr(number of recruits is in the jth recruit class given the stock size is in the j_sth stock class)

$$= \Pr(N_{ki} \in [(j-1)\hat{N}_k/m,\ j\hat{N}_k/m] \mid S_{i-k} \in [(j_s-1)\hat{S}/m_s,\ j_s\hat{S}/m_s]),$$

$$j_s = 1 \ldots m_s,\quad j = 1 \ldots m$$

The elements of T are estimated by the observed proportions of stock-recruit data points that lie in each of the (m_s x m) rectangles in the stock-recruit plane. Thus, no parametric model (such as those discussed in Section 2.3) is required to describe the stock-recruitment relationship. It is assumed that the only random variation in the population occurs in the stock-recruitment relationship.

The numbers of fish of age a ($a > k$) are also partitioned into age-classes. It is assumed that

$$N_{a+1,i+1} = N_{ai}\exp(-M_a - q_a e_i), \qquad a \geq k$$

Hence, the upper bound on the number of fish of age k+1 in year i+1 is

$$\hat{N}_{k+1,i+1} = \hat{N}_k \exp(-M_k - q_k e_i)$$

and if N_{ki} is in the jth subinterval of $[0, \hat{N}_k]$, then $N_{k+1,i+1}$ is in the jth subinterval of $[0, \hat{N}_{k+1,i+1}]$. Similarly, the upper bound on the number of fish of age k+2 in year i+2 is

$$\hat{N}_{k+2,i+2} = \hat{N}_{k+1,i+1} \exp(-M_{k+1} - q_{k+1}e_{i+1})$$

etc. (The interval $[0, \hat{N}_k]$ is the same for each year class; however, the intervals $[0, \hat{N}_{ai}]$ ($a > k$) depend on the year class because the fishing effort is year dependent).

Finally, probability matrices $Q(i)$ and $Q_s(i)$ are constructed, with elements

$Q(i)_{aj}$ = Pr(number of fish of age a in year i is in the jth subinterval)

$$= \Pr(N_{ai} \in [(j-1)\hat{N}_{ai}/m, j\hat{N}_{ai}/m]), \qquad a \geq k, \quad j = 1 \ldots m$$

$Q_s(i)_j$ = Pr(spawning stock in year i is the jth subinterval)

$$= \Pr(S_i \in [(j-1)\hat{S}/m_s, j\hat{S}/m_s]), \qquad j = 1 \ldots m_s$$

Having estimated T, the population can be projected into the future by a simulation. If the fishing effort is assumed constant, the probability matrices $\{Q(i)\}$ and $\{Q_s(i)\}$ reach an equilibrium position, and the means and variances of the yield and stock size can be obtained. Unfortunately, to calculate the matrices $\{Q(i)\}$ and $\{Q_s(i)\}$, it is necessary to assume that serial correlations between year classes are negligible; hence, the model is only useful when recruitment is loosely linked to stock level.

2.6 STATISTICAL PROBLEMS

The main problems in modelling the dynamics of real populations are caused by

i) the large random variation that occurs at many stages of the life cycle,

ii) the large measurement errors that arise in the estimation of population numbers, effort figures, etc, particularly for marine species.

The combination of these two types of "noise" often makes it very difficult to identify appropriate models with which to describe the population dynamics. For example, different models, such as the Ricker model (2.6a) and the Beverton and Holt model (2.6b), often fit stock-recruitment data equally well. Indeed, simulations have shown that i) and ii) above can lead to stock-recruitment data that visually suggest quite inappropriate stock-recruitment relationships (Walters and Ludwig, 1981). A large noise component in the data also increases the complexity of parameter estimation and inference. Sometimes, it is not even clear how many parameters are estimable; for example, for some parameter values, terms in Schnute's model (2.10) combine to leave other parameters confounded (Schnute 1985).

Measurement errors in population numbers and effort figures also lead to the classical "errors-in-variables" problem. Most fish population models are fitted by least squares regression in which the independent variates (generally stock sizes or effort figures) are assumed to be known without error. However, this assumption is rarely satisfied in practice. Hence, the least squares estimators are often biased and have a variance much greater than the minimum variance bound. For example, there is substantial bias in the estimators of the parameters of the Schaefer model (2.1 and 2.2a) obtained using any one of four different formulations of the model (eg (2.3) and (2.4)) (Uhler 1980).

A method of parameter estimation that accounts for the errors-in-variables in a stock-recruitment model is described by Ludwig and Walters (1981). They consider a seasonally exploited population with a one year life cycle in which the recruits and the spawning stock in year i are estimated with the same multiplicative measurement error. That is,

$$r_i = R_i \exp(\upsilon_i)$$

$$s_i = S_i \exp(\upsilon_i)$$

where

i) r_i (R_i) is the observed (true) number of recruits in year i,

ii) s_i (S_i) is the observed (true) spawning stock in year i,

iii) $\{\upsilon_i\}$ are NID$(0, \sigma_\upsilon^2)$.

The true number of recruits in year i+1 is related to the true spawning stock in year i by a Ricker stock-recruitment function (2.6a) with a multiplicative error

$$R_{t+1} = \alpha S_t \exp(-b S_t)\exp(\epsilon_t)$$

where $\{\varepsilon_i\}$ are NID(0, σ^2). Estimates of the stock-recruitment parameters and of the errors $\{\nu_i\}$ are found by maximum likelihood. Estimates of the true numbers of recruits and the true stock sizes can then also be calculated. However, the ratio σ_ν^2/σ^2 must be known a priori for the problem to be soluble; this quantity is difficult to obtain in practice, so the usefulness of the method is limited.

A further problem in least squares estimation is caused by the serial correlations that often exist between the dependent and independent variates (Walters, 1985). For example, consider a stock-recruitment model for a population with a one year life cycle, such that

$$R_{t+1} = f(S_t) + \epsilon_t,$$

$$S_{t+1} = \tau R_{t+1},$$

where $\{\varepsilon_i\}$ are NID(0, σ^2). The error term ε_i affects R_{i+1} and hence the independent variates S_{i+1}, S_{i+2}, etc; again, this leads to bias in the least squares estimators.

Many of the problems above are greatly reduced if there is a large contrast in the independent variates (ie the stock sizes and effort figures). This is because information is then available about the responses of the population to a wide variety of conditions (Ludwig and Walters, 1981; Ludwig and Hilborn, 1983). However, such contrast often does not exist, especially for populations which have only been studied since regular fishing has depleted the size of the stock. Ludwig and Hilborn

recommend that managers should actively generate contrast in effort, so that improved models can be developed and better long-term fishing policies can be formulated.

CHAPTER 3 MEASURES OF NONLINEARITY

3.1 INTRODUCTION

The theory of estimation and inference in linear least squares regression is well developed. For example, if the random errors are independent and normally distributed with zero mean and constant variance,

i) the distribution of the least squares estimator is known explicitly,

ii) the least squares estimator has the "optimum" properties that it is unbiased and normally distributed with a variance that attains the minimum variance bound,

iii) exact parameter confidence regions are easy to calculate and are conveniently expressed as ellipsoids in the parameter space.

Nonlinear least squares regression is more complicated and presents many problems of inference. For example, the distribution of the least squares estimator is usually not known explicitly. Also, exact parameter confidence regions must generally be calculated numerically and are difficult to display when there are more than two parameters. Hence, there is great value in using simple approximate inference methods that give adequate, easily presentable results. Inference is particularly simple if a nonlinear model can be approximated by a linear model; hence, most inference methods for nonlinear models are based on a linear approximation of the model about the least squares estimate (Chambers, 1973).

The linear approximation is obtained as follows. Consider the nonlinear regression model

$$y_i = f(x_i,\theta) + \epsilon_i, \quad i = 1...n \tag{3.1}$$

where

i) $\{y_i, i = 1...n\}$ is a set of observations related to a set of known design points $\{x_i, i = 1...n\}$ and an unknown p-vector of parameters $\theta = (\theta_1,...,\theta_p)'$ through a known model function $f(.)$,

ii) the additive random errors ϵ_i, $i = 1...n$, are NID$(0, \sigma^2)$.

If $y, \eta(\theta)$ and $\underset{\sim}{\epsilon}$ are the n-vectors

$$y = (y_1,...,y_n)',$$

$$\eta(\theta) = (f(x_1,\theta),...,f(x_n,\theta))',$$

$$\epsilon = (\epsilon_1,...,\epsilon_n)',$$

then (3.1) is conveniently expressed as

$$y = \eta(\theta) + \epsilon. \tag{3.2}$$

Let the least squares estimate of θ be $\hat{\theta}$, the value which minimises the residual sum of squares

$$S(\theta) = \sum_{i=1}^{n} (y_i - f(x_i,\theta))^2 = (y - \eta(\theta))'(y - \eta(\theta)). \tag{3.3}$$

Model (3.1) is linearised about $\hat{\theta}$ as

$$y_i \cong f(x_i,\theta) + \sum_{j=1}^{p} (\theta_j - \hat{\theta}_j) \left.\frac{df(x_i,\theta)}{d\theta_j}\right|_{\theta=\hat{\theta}} + \epsilon_i, \qquad i = 1 \ldots n,$$

or

$$y \cong \eta(\hat{\theta}) + V.(\theta - \hat{\theta}) \tag{3.4}$$

where $V.$ is the nxp matrix of first partial derivatives of f with ijth element

$$(V.)_{ij} = \left.\frac{df(x_i,\theta)}{d\theta_j}\right|_{\theta=\hat{\theta}}, \qquad i = 1 \ldots n, j = 1 \ldots p \tag{3.5}$$

Inferences are then made using standard linear theory. For example, $\hat{\theta}$ is approximately distributed $N(\theta, \sigma^2 (V.'V.)^{-1})$. Also, an approximate 95% confidence region for θ is given by

$$(\theta - \hat{\theta})'V.'V.(\theta - \hat{\theta}) \leq ps^2 F(p, \nu; 0.05), \tag{3.6}$$

where s^2 is an estimate of σ^2 based on ν degrees of freedom.

The validity of linear approximation inferences depends on the adequacy of the linear approximation. This can be assessed by a variety of measures of nonlinearity. If the nonlinearity of a model is small (or equivalently, if the nonlinear behaviour of the model is small or if the model is close-to-linear) then the linear approximation

is good and linear approximation inferences can be made with reasonable security. However, if the nonlinearity is large, then the linear approximation is poor and linear approximation inferences can give very misleading results.

Nonlinear regression models and measures of nonlinearity are used extensively throughout this Thesis. Hence, it is convenient to describe some standard terminology and the measures of nonlinearity in this Chapter. The solution locus and the curvature measures of Bates and Watts are described in Section 3.2. Measures of nonlinearity based on the distribution of the least squares estimator are described in Section 3.3. Finally, inference methods that can be used when the linear approximation is poor are discussed in Section 3.4.

3.2 THE CURVATURE MEASURES OF BATES AND WATTS

3.2.1 The Solution Locus

The parameters θ lie in a p-dimensional parameter space. Similarly, y lies in an n-dimensional sample space. The two spaces are "linked" by the function η through (3.2). As θ varies in the parameter space, $\eta(\theta)$ maps out a p-dimensional surface, called the solution locus, in the sample space (Box and Lucas, 1959). If there is no error (ie if $\varepsilon = 0$), then y lies in the solution locus; however, in general, y lies outside the solution locus. The least squares estimate $\hat{\theta}$ is the value of θ corresponding to the point in the solution locus $\hat{\eta} = \eta(\hat{\theta})$ closest to y. For example, consider the model

$$y_i = \frac{x_i}{1 + \theta x_i} + \epsilon_i, \qquad i = 1, 2, \tag{3.7a}$$

with $x_1 = 1.0$ and $x_2 = 2.0$. Suppose $y = (0.07, 0.11)'$ so that $\hat{\theta} = 10.3$. The solution locus is then the (1-dimensional) curve in the (2-dimensional) sample space consisting of all points of the form

$$\left(\frac{1}{1+\theta} \cdot \frac{2}{1+2\theta}\right)$$

(Fig. 3.1).

The parameter curves are the curves in the solution locus obtained by allowing one parameter, θ_i say, to vary whilst holding the others constant. In the example above, there is only one parameter, so the θ parameter curve is the solution locus. (However, it is also useful to consider the points in the solution locus corresponding to θ = constant as the (degenerate) parameter curves of a (degenerate) second parameter).

It is important to note that the solution locus is independent of the parameterisation of the model. However, different parameterisations can give markedly different forms of parameter curves in the solution locus. For example, the solution locus is unchanged if (3.7a) is reparameterised as

$$y_i = \frac{\phi x_i}{\phi + x_i} + \epsilon_i, \quad i = 1,2 \qquad (3.7b)$$

where $\phi = 1/\theta$ (Fig. 3.2). However, the spacing between the points in the solution locus corresponding to equispaced values of θ changes rapidly near $\hat{\eta}$, whereas the points corresponding to equispaced values of ϕ are themselves approximately equispaced near $\hat{\eta}$.

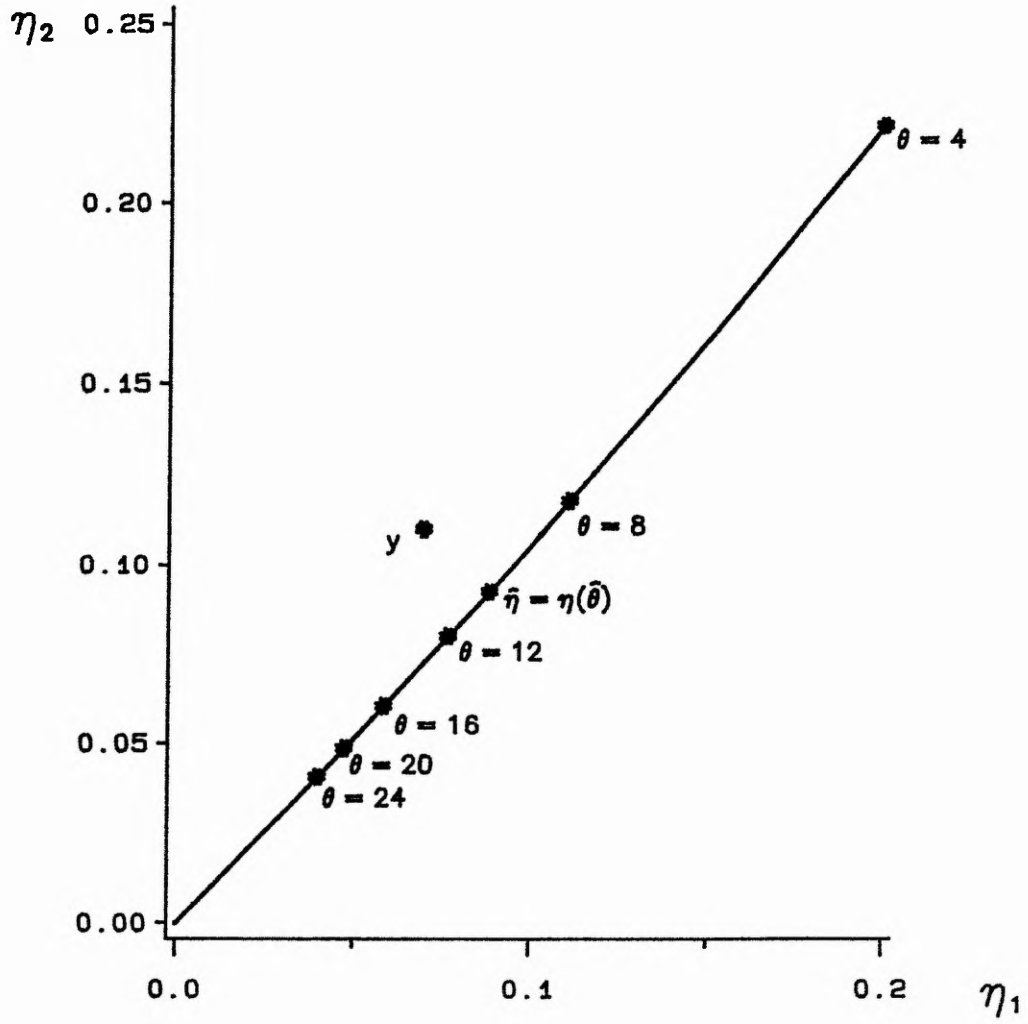


Fig. 3.1 The solution locus of model (3.7a).

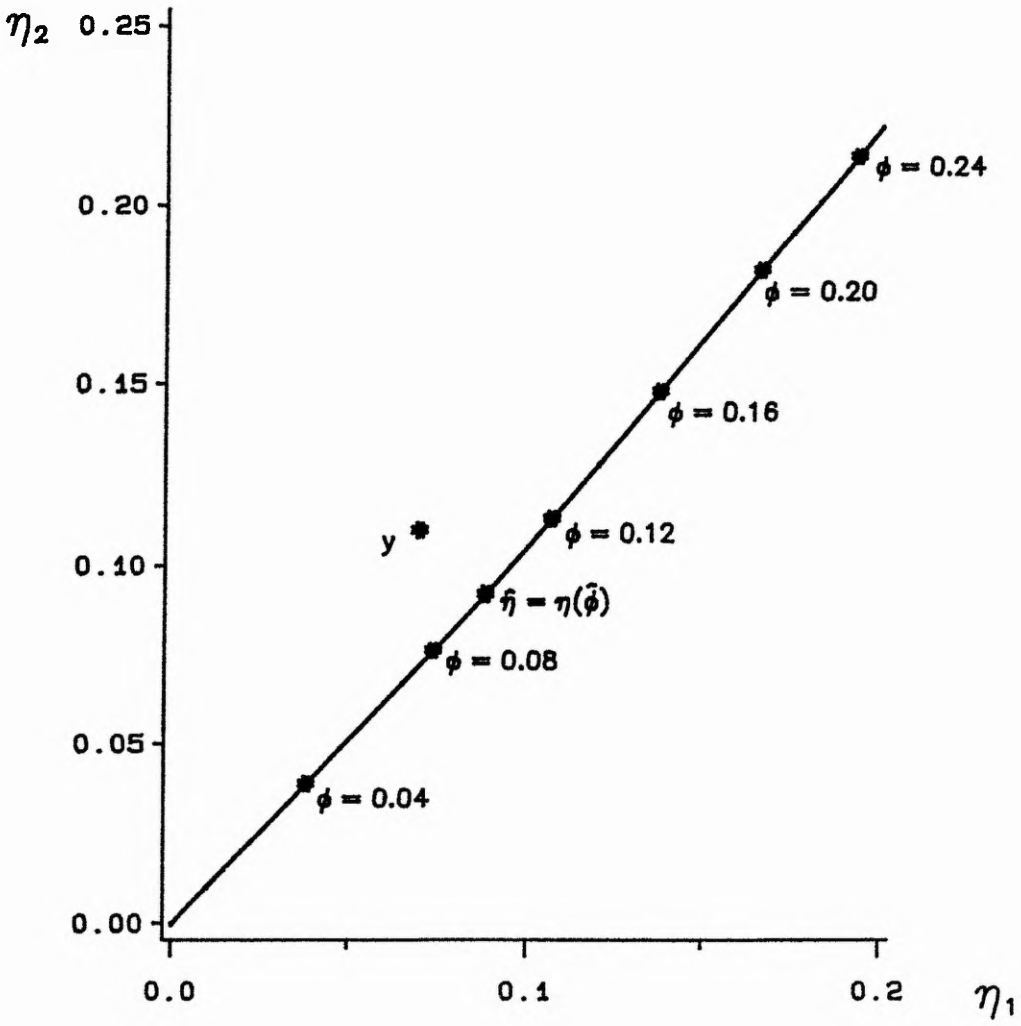


Fig. 3.2 The solution locus of model (3.7b).

3.2.2 Intrinsic and Parameter Effects Nonlinearity

The linearisation (3.4) consists of two distinct approximations:

i) the tangent plane approximation - the solution locus is approximated by its tangent plane at $\hat{\eta}$,

ii) the uniform co-ordinates approximation - the parameter curves in the solution locus are approximated by straight, parallel, equi-spaced lines in the tangent plane.

The extent to which the solution locus "deviates" from the tangent plane at $\hat{\eta}$ is independent of the parameterisation of the model (since the solution locus is independent of the parameterisation) and is called the intrinsic nonlinearity of the model (Bates and Watts, 1980). For example, the solution locus of model (3.7) is almost linear near $\hat{\eta}$ (Figs 3.1 and 3.2) so the tangent plane approximation is good and the intrinsic nonlinearity is small. (Strictly, the adequacy of the tangent plane approximation depends on a "region of interest" around $\hat{\eta}$. If inferences at the 100(1 - α)% confidence level are required, this region is taken to be the sphere in the tangent plane given by

$$(\eta - \hat{\eta})'(\eta - \hat{\eta}) \leq ps^2 F(p, \nu; \alpha),$$

since in the linear case a 100(1 - α)% parameter confidence region consists of all values Θ such that $\eta(\Theta)$ lies in this sphere. The intrinsic nonlinearity is said to be small in the 100(1 - α)% confidence region if the solution locus is well approximated by its tangent plane at $\hat{\eta}$ in the 100(1 - α)% confidence region of interest).

The extent to which the parameter curves "deviate" from straight, parallel, equispaced lines is called the parameter effects nonlinearity (Bates and Watts, 1980). This does depend on the model parameterisation. For example, parameterisation (3.7a) has large parameter effects nonlinearity because the spacings between the θ parameter curves change rapidly near $\hat{\eta}$ (Fig. 3.1). However, parameterisation (3.7b) has small parameter effects nonlinearity because the ϕ parameter curves are almost equispaced (Fig. 3.2).

If the intrinsic nonlinearity is small, inferences based on the tangent plane approximation can be made with security; for example, standard F tests can be used to test hypotheses about subsets of the parameter values, confidence intervals can be constructed for predicted values of y, etc. If both types of nonlinearity are low, then any standard linear theory can be used with security. All linear approximation inferences are likely to be invalid if the intrinsic nonlinearity is large, regardless of the size of the parameter effects nonlinearity. However, if the intrinsic nonlinearity is small but the parameter effects nonlinearity is large, alternative parameterisations can be investigated to reduce the parameter effects nonlinearity.

3.2.3 The Bates and Watts Curvature Measures of Nonlinearity

Bates and Watts (1980) developed measures of intrinsic and parameter effects nonlinearity based on the curvature of the solution locus and the parameter curves at $\hat{\eta}$. First, consider the intrinsic nonlinearity. This is small if the solution locus is approximately flat over the region of interest; ie if the curvature of the solution locus at $\hat{\eta}$ is small. Since the solution locus is a p-dimensional surface, its curvature at $\hat{\eta}$ varies with the "direction of travel through $\hat{\eta}$". For example, suppose the solution locus consists of a long thin valley; the curvature of the solution locus is

much greater travelling from one side of the valley to the other than travelling along the floor of the valley. Bates and Watts construct measures of intrinsic curvature that correspond to each direction of travel through $\hat{\eta}$ and hence define the maximum intrinsic curvature Γ^N (corresponding to the maximum curvature over all directions) and the root mean square (RMS) intrinsic curvature $\gamma^N{}_{RMS}$ (corresponding to an average curvature). Values of Γ^N and $\gamma^N{}_{RMS}$ less than $1/(2\sqrt{F(p,\upsilon;\alpha)})$ indicate that the solution locus is well approximated by the tangent plane over the $100(1-\alpha)\%$ region of interest. For example, suppose that $\sigma^2 = 0.001$ in model (3.7); then

$$\Gamma^N = \gamma^N_{RMS} = 0.011,$$

(the two measures are identical for a 1-parameter model); the curvature measures are much less than $1/(2\sqrt{F(1,\infty;0.05)}) = 0.26$, so the intrinsic nonlinearity of the model is small in the 95% confidence region.

The maximum parameter effects curvature Γ^T and the RMS parameter effects curvature $\gamma^T{}_{RMS}$ are developed in a similar way to the intrinsic curvature measures by considering the curvature of the parameter curves at $\hat{\eta}$. Values of Γ^T and $\gamma^T{}_{RMS}$ greater than $1/(2\sqrt{F(p,\upsilon;\alpha)})$ indicate large parameter effects nonlinearity. For parameterisation (3.7a),

$$\Gamma^T = \gamma^T_{RMS} = 0.49,$$

which is large compared to $1/(2\sqrt{F}) = 0.26$ so the parameter effects nonlinearity is large. However, for parameterisation (3.7b),

$$\Gamma^T = \gamma^T_{RMS} = 0.035,$$

which is small compared to $1/(2\sqrt{F})$, so the ø parameter curves in the region of interest are well approximated by straight, parallel, equispaced lines.

There is some dispute over whether the maximum or the RMS curvature measures should be used to assess nonlinearity (since when $p > 1$, Γ^N can be greater than $1/(2\sqrt{F})$ whilst at the same time γ^N_{RMS} is less than $1/(2\sqrt{F})$ etc) (see the discussion of Bates and Watts (1980)). However, the "critical" value $1/(2\sqrt{F})$ is chosen because it is a sensible value which the curvature measures can be compared to and not because it represents any formal "significance test of nonlinearity" (Healy, 1984). Hence, a sensible approach is to calculate both sets of measures and use these in conjunction with the measures described in Sections 3.2.4 and 3.3 to obtain an overall picture of the nonlinearity of the model.

The Bates and Watts curvature measures are calculated using the first and second derivatives of the model function at $\hat{\theta}$. Essentially, these derivatives are used to construct a quadratic approximation to the solution locus and the parameter curves over the region of interest. This approach is satisfactory for the vast majority of nonlinear models. However, in Chapter 7, it is shown that the Bates and Watts curvature measures can not be applied to nonlinear models with change points. These models have discontinuous derivatives, so the form of the solution locus and the parameter curves can change markedly within the region of interest. In this case, the quadratic approximation breaks down and the curvature measures can seriously underestimate the nonlinearity of the model.

3.2.4 Axis Ratio Measures of Intrinsic Nonlinearity

Parameter inference regions (ie likelihood and confidence regions) can be thought of as the images in parameter space of inference regions in the solution locus. For example, likelihood regions consist of all points θ in the parameter space such that

$$S(\theta) - S(\hat{\theta}) < \delta^2.$$

where $S(\theta)$ is given by (3.3) and δ^2 defines the level of the likelihood region. These regions correspond to likelihood regions in the solution locus consisting of all points $\eta = \eta(\theta)$ such that

$$(y-\eta)'(y-\eta) < S(\hat{\theta}) + \delta^2$$

The axis ratio measures of intrinsic nonlinearity are based on a comparison of linear and quadratic approximations to solution locus inference regions (Hamilton, Watts and Bates, 1982). In a linear approximation, solution locus inference regions are approximated by spheres in the tangent plane; the adequacy of this approximation depends on the intrinsic nonlinearity. (Parameter inference regions are then obtained by approximating the images of these spheres by ellipsoids in parameter space). However, in a quadratic approximation, solution locus inference regions correspond to ellipsoids in the tangent plane, rather than spheres. Hence, measures of intrinsic nonlinearity can be based on the differences between the quadratic approximation ellipsoids and the linear approximation spheres. The axis ratio measures are defined as the ratios of the maximum and minimum axis lengths of the ellipsoid to the radius of the sphere. These are calculated for both likelihood and confidence regions; values close to one (ie within 0.1 of unity) indicate small intrinsic nonlinearity. In

model (3.7), the 95% confidence region axis ratio is 1.009 and the corresponding likelihood axis ratio is 1.004; both indicate small intrinsic nonlinearity. (There is only one axis ratio because $p = 1$.)

The quadratic approximation ellipsoids can also be used to compensate for the effect of intrinsic nonlinearity on parameter inference regions. The ellipsoid is approximated by a sphere with radius given by the largest axis length of the ellipse. The image in parameter space of this "inflated" sphere is then a conservative inference region for θ. A measure of the adequacy of the quadratic approximation is given by Hamilton, Watts and Bates (1982). This is expressed as a percentage; values greater than 5% indicate that the quadratic approximation is poor and that conservative inference regions do not well account for intrinsic nonlinearity. In model (3.7), the quadratic approximation measure is 0.01% indicating that the approximation is very good.

The conservative inference regions described above do not compensate for parameter effects nonlinearity. Suitable ways of reparameterising models to reduce large parameter effects nonlinearity are discussed in Bates and Watts (1981).

3.3 NONLINEARITY MEASURES BASED ON THE DISTRIBUTION OF $\hat{\theta}$

In a linear model, the least squares estimator is unbiased and normally distributed with a variance that attains the minimum variance bound. In a nonlinear model, this is true only asymptotically (ie as $n \to \infty$); for small samples, $\hat{\theta}$ can be biased, skewed and have a variance much greater than the minimum variance bound. This Section describes measures of nonlinearity that compare the distribution of $\hat{\theta}$ with

the distribution of the least squares estimator of a linear model; a nonlinear model is close-to-linear if $\hat{\theta}$ is approximately unbiased and approximately normally distributed with a variance close to the minimum variance bound.

Nonlinearity measures based on the distribution of $\hat{\theta}$ do not distinguish between the effects of intrinsic and parameter effects nonlinearity. However, they are generally applied to each parameter $\theta_1,...,\theta_p$ in turn and indicate which parameters have close-to-linear behaviour and which do not. Hence, they are particularly useful for indicating which parameters should be reparameterised to reduce large parameter effects nonlinearity.

3.3.1 Simulation Approach

The distribution of $\hat{\theta}$ can be investigated by simulations, as described in great detail by Ratkowsky (1983). It is assumed that the true values of θ and σ^2 are given by $\hat{\theta}$ and s^2 respectively. These values are used, in conjunction with (3.1), to generate a large number N (1000, say) of simulated data sets, from which a set of least squares estimates $\{\tilde{\theta}_j = (\tilde{\theta}_{1j},...,\tilde{\theta}_{pj})', j = 1...N\}$ are obtained. Histograms of the simulated estimates are often sufficient to indicate large nonlinear behaviour. For example, the simulated distribution of $\hat{\theta}$ in model (3.7a) is highly skewed (Fig. 3.3).

Quantitative measures of nonlinearity are obtained by calculating the first four sample moments of the simulated distribution of each parameter θ_i, i = 1...p; namely

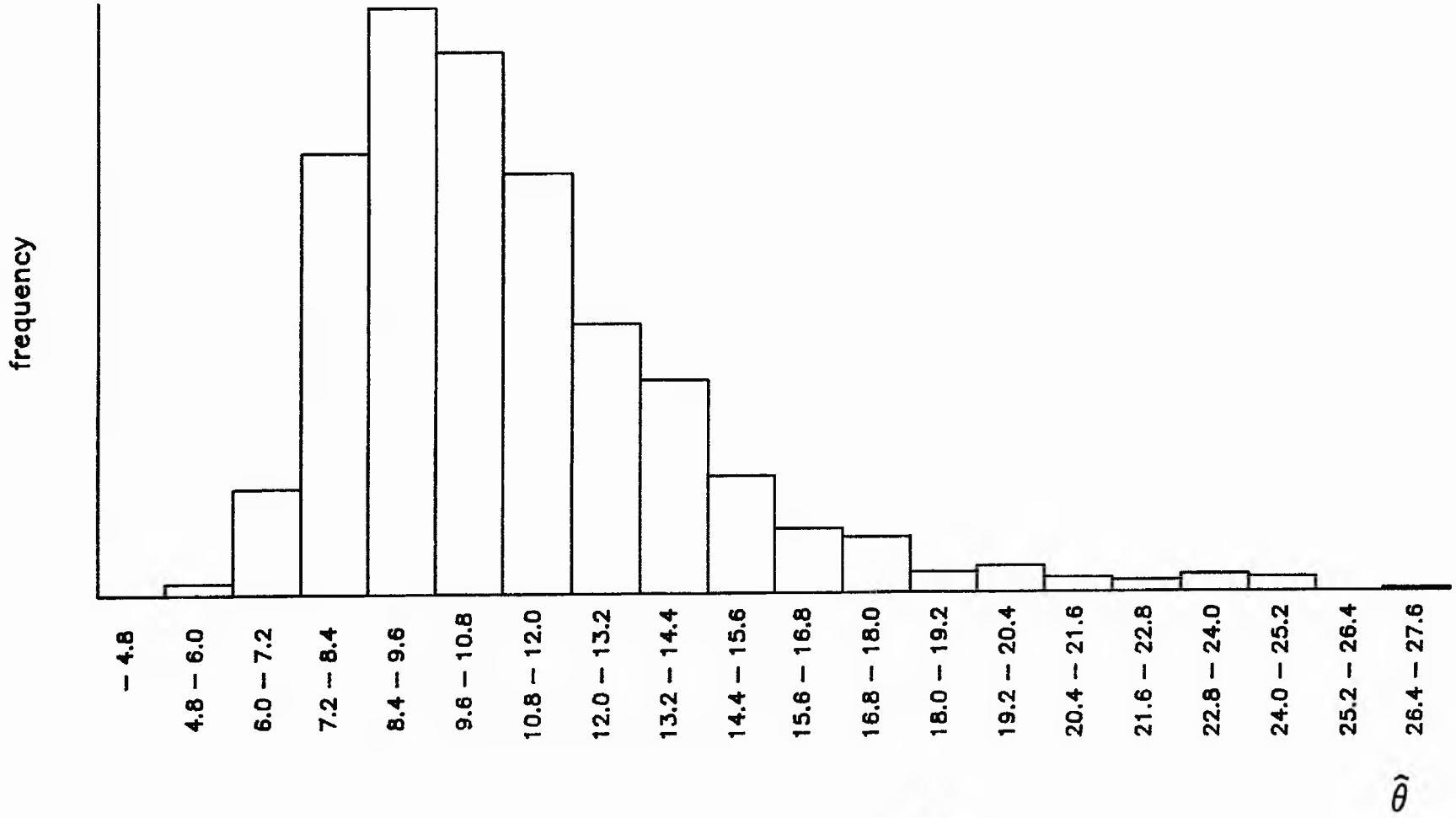


Fig. 3.3 A histogram of the simulated values of $\hat{\theta}$ in model (3.7a).

$$m_{i1} = \frac{1}{N} \sum_{k=1}^{N} \hat{\theta}_{ik}$$

$$m_{ij} = \frac{1}{N} \sum_{k=1}^{N} (\hat{\theta}_{ik} - m_{i1})^j, \quad j = 2,3,4$$

Measures of the bias of $\hat{\theta}_i$, the variance of $\hat{\theta}_i$ relative to the minimum variance bound, the skewness of $\hat{\theta}_i$ and the excess kurtosis of $\hat{\theta}_i$ are then given by T_{i1}, T_{i2}, T_{i3}, T_{i4} respectively, where

$$T_{i1} = (m_{i1} - \theta_i)/\sqrt{(\tau_i^2/N)}$$

$$T_{i2} = \sqrt{2(N-1)\, m_{i2}/\tau_i^2} - \sqrt{2N-3}$$

$$T_{i3} = \frac{m_{i3}}{(m_{i2})^{3/2}} \sqrt{\frac{N}{6}}$$

$$T_{i4} = \left(\frac{m_{i4}}{(m_{i2})^2} - 3 \right) \sqrt{\frac{N}{24}}$$

and where τ_i^2 is the minimum variance bound of $\hat{\theta}_i$ (estimated by the ith diagonal element of $s^2 (V.'V.)^{-1}$ (from (3.5))). If $\hat{\theta}_i$ has close-to-linear behaviour, then $\hat{\theta}_i$ is approximately distributed $N(\theta_i, \tau_i^2)$ and for large N, $\{T_{ij}\}$ are all approximately distributed N(0,1). Thus, the nonlinear behaviour of $\hat{\theta}_i$ is assessed by comparing T_{i1}, T_{i2}, T_{i3}, T_{i4} to a N(0,1) distribution.

For parameterisation (3.7a),

$$T_1 = 9.79, \quad T_2 = 14.70, \quad T_3 = 32.08, \quad T_4 = 73.77,$$

which are all very large compared to a N(0,1) distribution. Hence, $\hat{\theta}$ is biased, skewed and has a variance much greater than the minimum variance bound and a kurtosis much greater than the kurtosis of a normal distribution. However, for parameterisation (3.7b),

$$T_1 = -0.34, \quad T_2 = -1.90, \quad T_3 = 1.10, \quad T_4 = 1.09,$$

which are all within the interval (-1.96, 1.96) (ie the standard 95% interval for a N(0,1) random variable); hence, model (3.7b) is close-to-linear.

3.3.2 Asymmetry Measures of Nonlinearity

In many cases, large nonlinear behaviour is characterised by asymmetry in the distribution of $\hat{\theta}$; ie by bias and skewness (Lowry and Morton, 1983). Hence, measures of asymmetry are used to assess the nonlinearity of a model.

The asymmetry measure of Lowry and Morton (1983) is based on a comparison of the least squares estimator $\hat{\theta}_+$ when

$$y = \eta(\theta) + \epsilon,$$

with the corresponding least squares estimator $\hat{\theta}_-$ when

$$y = \eta(\theta) - \epsilon$$

(ie when the sign of the error term is changed). In a linear model, $\hat{\theta}_+ + \hat{\theta}_- = \text{constant}$. However, in a nonlinear model, this is generally not the case. If

$$\psi = (\hat{\theta}_+ + \hat{\theta}_-)/2,$$

then

$$\lambda_j = \mathrm{Var}(\psi_j)/\mathrm{Var}(\hat{\theta}_j)$$

is a measure of the asymmetry of the distribution of $\hat{\theta}_j$, $j = 1...p$. If the distribution of $\hat{\theta}_j$ is symmetric, then $\lambda_j = 0$; as the asymmetry of the distribution increases, λ_j increases towards 1. Values of $\lambda_j > 0.01$ indicate a degree of asymmetry; if $\lambda_j > 0.05$, then asymmetry is generally clear from a histogram of simulated values of $\hat{\theta}_j$. The measures (λ_j) are calculated either by a simulation (ie by generating a large number of errors ϵ and then estimating $\hat{\theta}_+$ and $\hat{\theta}_-$ for each error) or by an asymptotic formula based on the second derivatives of the model function at $\hat{\theta}$.

Two further measures of asymmetry are

i) Box's expression for the bias of $\hat{\theta}$ (Box, 1971); this is particularly useful for comparing the biases of different parameterisations,

ii) Hougaard's expression for the skewness of $\hat{\theta}$ (Hougaard, 1985); values of Hougaard's skewness greater than 0.2 indicate a skewed distribution.

Both these expressions are calculated using the first and second derivatives of the model function at $\hat{\theta}$.

For example, the asymmetry measures are

$\lambda = 0.198$, (simulated),

$\lambda = 0.121$, (asymptotic formula),

Box's bias = 0.667,

Hougaard's skewness = 1.48,

for parameterisation (3.7a) and

$$\lambda = 0.0007, \quad \text{(simulated)},$$

$$\lambda = 0.0007, \quad \text{(asymptotic formula)},$$

$$\text{Box's bias} = 0.0004,$$

$$\text{Hougaard's skewness} = 0.10,$$

for parameterisation (3.7b). Again, these measures indicate that model (3.7a) has large nonlinear behaviour and that model (3.7b) is close-to-linear.

3.4 LARGE NONLINEAR BEHAVIOUR

If the nonlinearity of a model is large (and can not be reduced by reparameterisation), then linear approximation inference methods are likely to give very misleading results. However, the only general alternative appears to be a simulation based approach. This can be used in a wide variety of situations. For example, approximate confidence intervals for (θ_i) can be obtained from the simulated distribution of $\hat{\theta}$ obtained in Section 3.3.1. Also, suppose it is required to test the hypothesis $\theta_1 = \theta_2$ against the hypothesis $\theta_1 \neq \theta_2$. Standard methodology involves fitting model (3.1) with

i) the constraint $\theta_1 = \theta_2$, giving residual sum of squares Q_1,

ii) θ_1, θ_2 unconstrained, giving residual sum of squares Q_0.

The test statistic

$$F = (n - p) \frac{(Q_1 - Q_0)}{Q_0}$$

is then calculated. If the intrinsic nonlinearity is small, then under the null hypothesis, F has an approximate central $F_{1,n-p}$ distribution. If the intrinsic nonlinearity is large, the distribution of F is not known. However approximate percentage points can be calculated by assuming that $\theta_1 = \theta_2$ and are given by the least squares estimates obtained in i) above, simulating a number of data sets *y* and then calculating F for each data set. Simulation methods are very versatile. Unfortunately, they also tend to be time consuming and to require a lot of expensive computer time, especially when *p* is large.

CHAPTER 4 NONLINEAR BEHAVIOUR IN THE ADDITIVE ERROR RICKER MODEL

4.1 INTRODUCTION

The Ricker model (2.6a) is the best general stock-recruitment model for describing the dynamics of the Black Brows Beck population (Elliott, 1985b). Two forms of the model are particularly useful:

i) the additive error Ricker model

$$R_i = a_1 E_i \exp(-b_1 E_i) + \epsilon_i, \quad i = 67...84, \tag{4.1}$$

where $\{\epsilon_i\}$ are NID$(0,\sigma^2)$,

ii) the multiplicative error Ricker model

$$R_i = a_1 E_i \exp(-b_1 E_i)\exp(\epsilon_i), \quad i = 67...84, \tag{4.2}$$

where $\{\epsilon_i\}$ are NID$(0,\sigma^2)$.

In (4.1) and (4.2), E is the number of eggs and R is the number of recruits at a given stage of the life cycle. The additive error model (4.1) is fitted by nonlinear least squares regression. The multiplicative error model (4.2) is fitted by linearising (4.2) as

$$\log(R_i) = \log(E_i) + a_2 - b_1 E_i + \epsilon_i, \quad i = 67...84, \tag{4.3}$$

where $a_2 = \log(a_1)$ and then using linear least squares regression.

Models (4.1) and (4.2) are both excellent descriptions of the relationship between R_1, the number of 0+ parr at time t_1 (the May/June electro-fishing) and E. The fitted curves are shown in Figures 4.1 and 4.2. An examination of the residuals gives no evidence that either of the error assumptions are unreasonable. In particular, there is no evidence that the variance of the residuals from the additive error model increases with the fitted value, which would indicate that the multiplicative error model is more appropriate.

Inference methods for the multiplicative error model are straightforward because (4.3) is a linear model. The least squares estimates and exact 95% confidence limits for a_2 and b_1 are given in Table 4.1; also given are the least squares estimate and exact 95% confidence limits for a_1 which are calculated from the relationship $a_1 = \exp(a_2)$.

Inference is more complicated for the additive error model (4.1) because it is nonlinear (Chapter 3). The least squares estimates and linear approximation 95% confidence limits for a_1, b_1 (and hence a_2) are also given in Table 4.1. However, the adequacy of these linear approximation confidence limits (and of other inferences based on a linear approximation) depends on the nonlinear behaviour of the model. Hence, in this chapter, the nonlinearity measures described in Chapter 3 are used to assess the nonlinear behaviour of the additive error Ricker model (4.1).

The nonlinearity of a model depends on both the parameterisation of the model and the data set that the model is fitted to. The nonlinearity of model (4.1) in combination

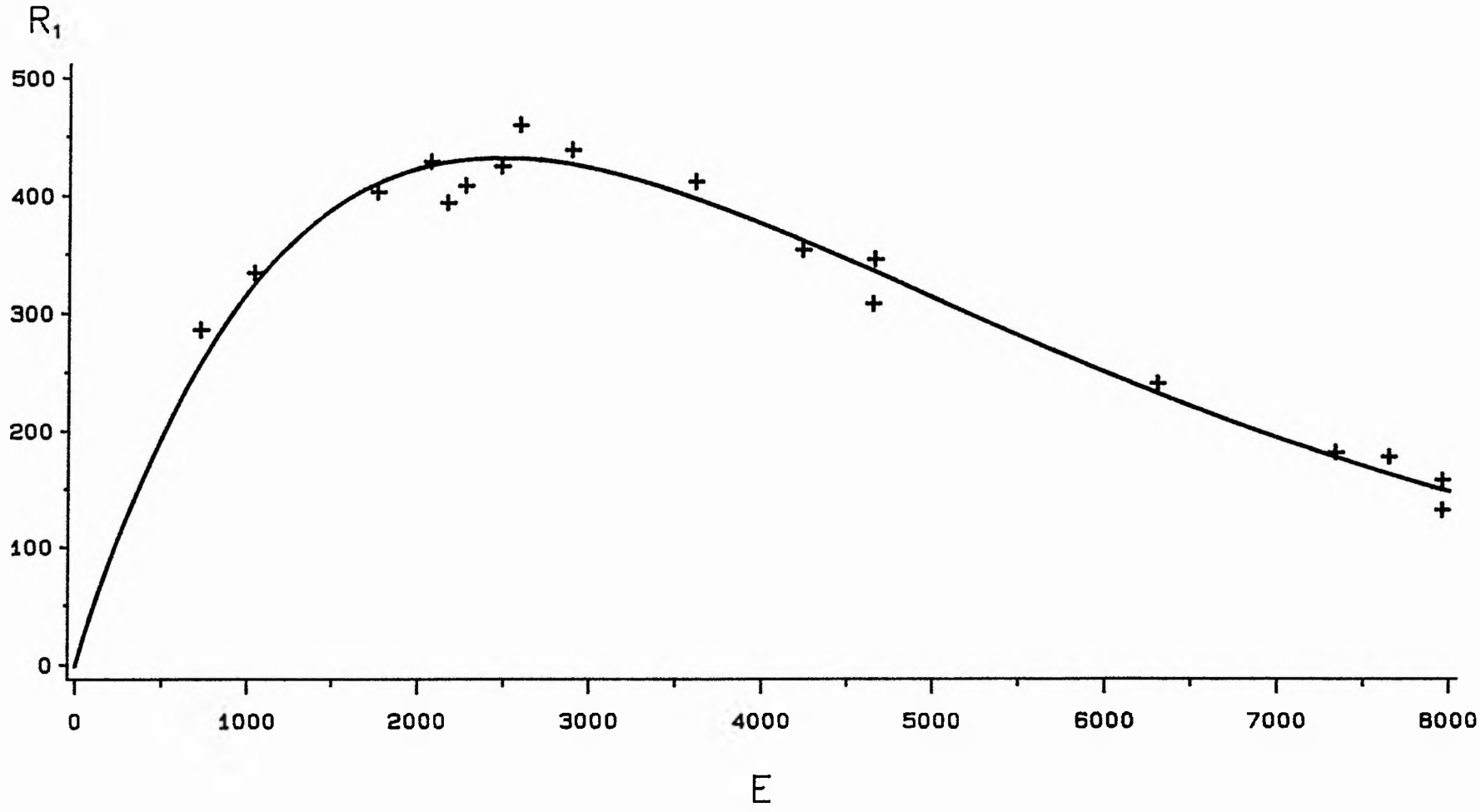


Fig. 4.1 The additive error Ricker model (4.1) fitted to the R_1 data.

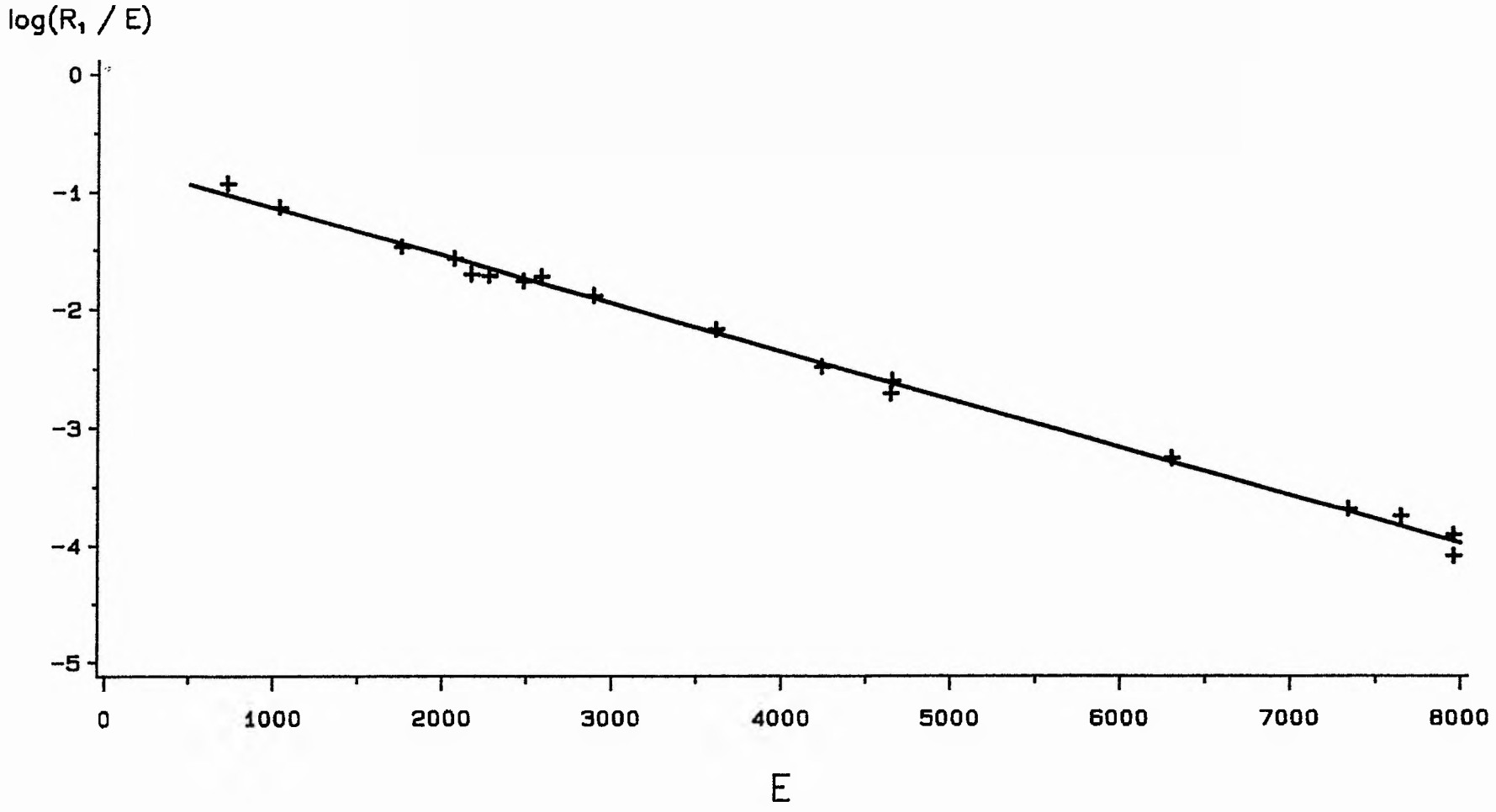


Fig. 4.2 The multiplicative error Ricker model (4.3) fitted to the R_1 data.

TABLE 4.1

Parameter estimates for models (4.1) and (4.3)

Parameter	Additive error model (4.1)	Multiplicative error model (4.2)
a_1	0.4775 (0.4491 0.5059)	0.4823 (0.4531 0.5134)
a_2	-0.7392 (-0.8005 -0.6814)	-0.7291 (-0.7916 -0.6666)
b_1	$4.065*10^{-4}$ ($3.891*10^{-4}$ $4.239*10^{-4}$)	$4.076*10^{-4}$ ($3.942*10^{-4}$ $4.210*10^{-4}$)
σ^2	347.2	$4.085*10^{-3}$

(95% confidence limits in brackets)

with the R_1 data set is of particular interest. (Strictly, the Ricker models do not adequately describe the relationships between the other life stages R_2, R_3, R_4, E^* and E because they fail to account for the effect of summer drought on survival; stock-recruitment models for these life stages are developed in Chapter 6). However, for some generality, the nonlinear behaviour of model (4.1) in combination with two other data sets is also investigated. These data sets are:

i) R_1' - a data set generated by model (4.1) with $a_1 = 0.4775$, $b_1 = 4.065*10^{-4}$ (the least squares estimates from the R_1 data set) and $\sigma = 50.0$,

ii) R_2' - the R_2 data set with the points from the years 1983 and 1984 (in which drought had a serious effect on survival) omitted.

The R_1' data set is generated with a large value of σ to give a model, data set combination with relatively large nonlinear behaviour. The R_2' data set is chosen because the least squares estimates obtained from R_2' are very different from those obtained from R_1 and R_1'. The three data sets are given in Table 4.2. (Data sets with markedly different sets of E values are not investigated).

4.2 MEASURES OF NONLINEARITY

The least squares estimates of a_1 and b_1 and the measures of nonlinearity are given in Table 4.3 for the three data sets R_1, R_1' and R_2'.

The intrinsic nonlinearity is small in each case. The Bates and Watts intrinsic curvatures are small, so the solution loci are well approximated by the tangent planes in the regions of interest. Also, the axis ratios are close to one so intrinsic nonlinearity has only a small effect on likelihood and 95% confidence regions.

TABLE 4.2

The R_1, R_1' and R_2' data sets

Year	E	R_1	R_1'	R_2'
1967	4652	346	340	116
1968	2272	409	435	92
1969	1756	403	389	74
1970	2584	460	360	115
1971	4644	308	374	110
1972	2888	439	477	112
1973	3610	412	511	119
1974	7958	132	207	74
1975	7334	182	145	88
1976	4234	354	437	91
1977	6300	241	166	99
1978	2068	429	351	105
1979	1034	334	396	73
1980	722	286	257	61
1981	7646	178	182	87
1982	7958	158	127	81
1983	2478	425	445	-
1984	2166	394	515	-

TABLE 4.3

Measures of nonlinearity for the R_1, R_1' and R_2' data sets

Data set	R_1 (May/June)	R_1' (May/June)	R_2' (Aug/Sep)
Intrinsic curvature:			
max	0.025	0.074	0.034
RMS	0.015	0.045	0.021
Axis ratios:			
likelihood	1.00 1.01	0.98 1.00	1.00 1.01
95% confidence	1.00 1.02	0.96 1.00	1.00 1.03
Quadratic error	0.05%	0.41%	0.09%
Parameter effects curvature:			
max	0.068	0.199	0.146
RMS	0.044	0.129	0.089
$1/(2\sqrt{F(p,n-p;0.05)})$	0.26	0.26	0.26
Parameter: a_1			
Least squares estimate	0.4775	0.4939	0.08189
Lin approx s.e.	0.0133	0.0404	0.00467
Box's bias	0.002	0.0019	0.00011
Hougaard's skewness	0.075	0.219 *	0.110
Lowry and Morton's λ	0.001	0.006	0.001
Parameter b_1			
Least squares estimate	$4.065*10^{-4}$	$4.036*10^{-4}$	$2.658*10^{-4}$
Lin approx s.e.	$0.082*10^{-4}$	$0.238*10^{-4}$	$0.122*10^{-4}$
Box's bias	$0.001*10^{-4}$	$0.006*10^{-4}$	$0.001*10^{-4}$
Hougaard's skewness	0.053	0.153	0.038
Lowry and Morton's λ	0.001	0.005	0.001

* indicates moderate nonlinear behaviour
** indicates serious nonlinear behaviour

The parameter effects nonlinearity is small for the R_1 data set; the Bates and Watts parameter effects curvatures are small, so the uniform co-ordinates approximation is acceptable in the region of interest. Also, the asymmetry measures indicate that $\hat{a}_1$ and $\hat{b}_1$ have close-to-linear behaviour. The parameter effects nonlinearity is greater for the other two data sets, but it should still not have a serious effect on parameter inference. For example, a conservative linear approximation 95% confidence region for a_1 and b_1 for the R_1' data set is close to an exact 95% confidence region (Fig. 4.3). However, the maximum parameter effects curvature for the R_1' data set is close to the "critical" value $1/(2\sqrt{F})$ and Hougaard's skewness indicates that the distribution of $\hat{a}_1$ is skewed. Hence, a number of alternative parameterisations are investigated to reduce the parameter effects nonlinearity for the R_1' data set; the effect of the reparameterisations on the other two data sets is also investigated.

4.3 REPARAMETERISATIONS

The parameter effects nonlinearity of the parameterisation

$$R_i = a_1 E_i \exp(-b_1 E_i) + \epsilon_i \qquad (4.4)$$

is sufficiently small for each of the three data sets, for the uniform co-ordinates approximation to be acceptable in each case. In addition, the parameters a_1 and b_1 have the important interpretations of measures of density independent and density dependent survival respectively (Section 2.3). Thus, a reparameterisation will be an improvement on (4.4) only if

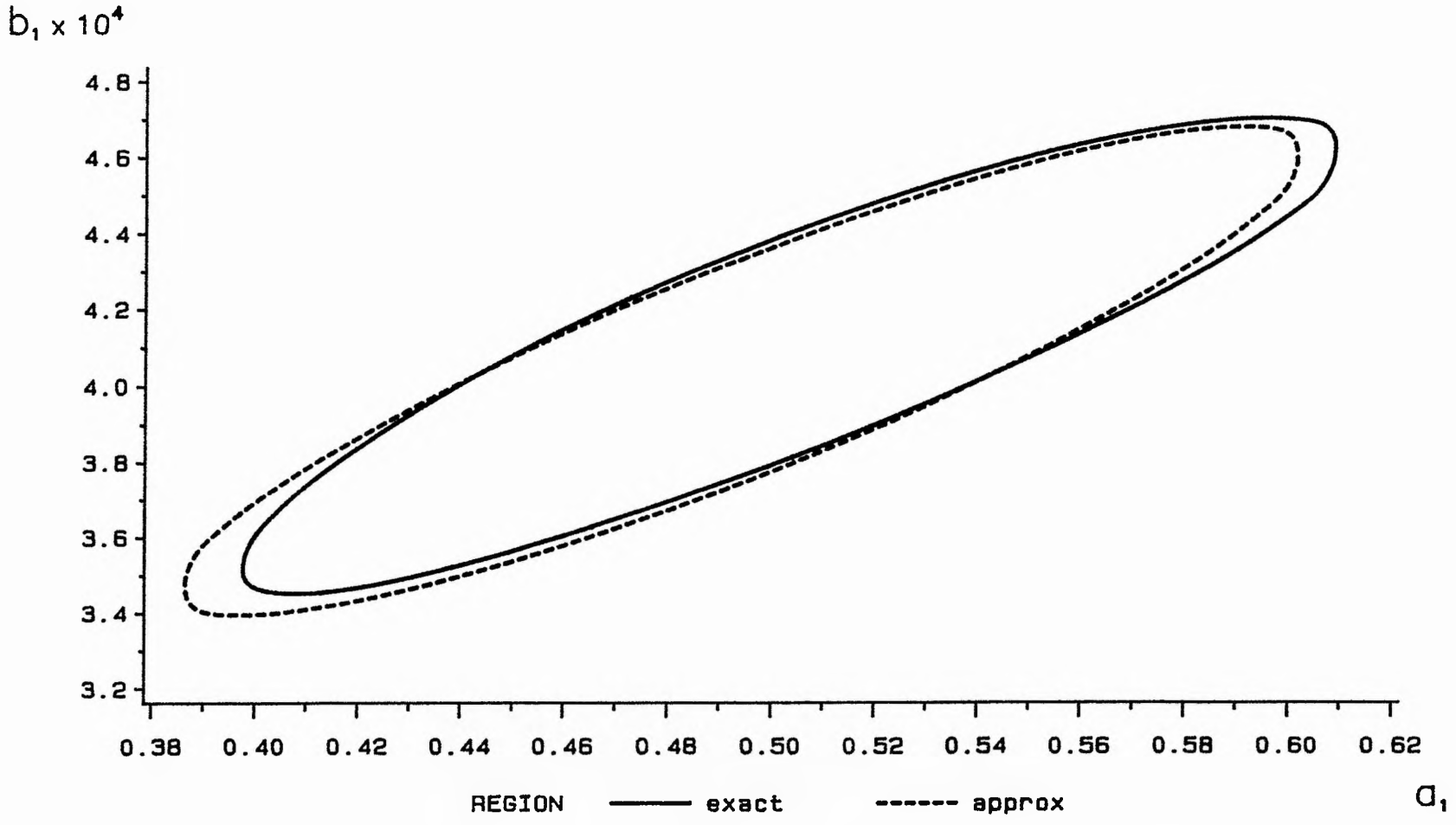


Fig. 4.3 Exact and linear approximation 95% confidence regions for a_1 and b_1 (R_1' data set).

i) the parameter effects nonlinearity is smaller than that of parameterisation (4.4),

ii) the new parameters are easily interpretable; ideally they should be measures of density independent and density dependent survival.

A parameterisation with very low parameter effects nonlinearity, but with parameters that are not easy to interpret, is not satisfactory. This is because the benefits of a slightly improved linear approximation are outweighed by the disadvantages of having to manipulate the new model to obtain information about the parameters of interest a_1 and b_1.

Parameters are easily interpretable if they are monotonic transformations of a_1 or b_1. For example, suppose the Ricker model is reparameterised as

$$R_t = a_1 E_t \exp(-\exp(b_2) E_t) + \epsilon_t$$

The parameter b_2 is a measure of density dependent survival since $b_2 = \log(b_1)$ is a monotonic transformation of b_1. When b_2 is very large and negative, survival is virtually density independent. As b_2 increases, survival becomes more density dependent. Thus, to compare the density dependence at different stages of the life cycle, it is sufficient to compare the corresponding values of b_2.

Only reparameterisations that involve monotonic (or other very simple) transformations of a_1 and b_1 are investigated; these include a number of recommended transformations for models with an exponential term (Bates and Watts, 1981; Draper and Smith, 1981). The transformations are

i) taking the logarithm of the scalar parameter a_1

$$R = E\exp(\alpha_2 - b_1 E) \qquad \alpha_2 = \log(\alpha_1)$$

ii) subtracting the average number of eggs $\bar{E}$ from the exponent

$$R = \alpha_3 E\exp(-b_1(E-\bar{E})) \qquad \alpha_3 = \alpha_1\exp(-b_1\bar{E}),$$

iii) taking the logarithm of the scalar parameter a_3

$$R = E\exp(\alpha_4 - b_1(E-\bar{E})) \qquad \alpha_4 = \log(\alpha_1) - b_1\bar{E}$$

iv) taking the logarithm of the exponential parameter b_1

$$R = \alpha_1 E\exp(-\exp(b_2)E) \qquad b_2 = \log(b_1),$$

v) taking the exponential of the exponent parameter b_1

$$R = \alpha_1 E(b_3)^{-E/1000} \qquad b_3 = \exp(1000b_1)$$

(where the factor of 1000 is introduced for numerical stability).

These transformations are combined to give the twelve parameterisations P1, P2,..., P12 listed in Table 4.4. The Bates and Watts parameter effects curvatures for each

TABLE 4.4

Model parameterisations

P1	$R = a_1 E \exp(-b_1 E)$	
P2	$R = E \exp(a_2 - b_1 E)$	$a_2 = \log(a_1)$
P3	$R = a_3 E \exp(-b_1 (E - \bar{E}))$	$a_3 = a_1 \exp(- b_1 \bar{E})$
P4	$R = E \exp(a_4 - b_1 (E - \bar{E}))$	$a_4 = \log(a_1) - b_1 \bar{E}$
P5	$R = a_1 E \exp(-\exp(b_2) E)$	$b_2 = \log(b_1)$
P6	$R = E \exp(a_2 - \exp(b_2) E)$	$a_2 = \log(a_1)$ $b_2 = \log(b_1)$
P7	$R = a_3 E \exp(- \exp(b_2)(E - \bar{E}))$	$a_3 = a_1 \exp(- b_1 \bar{E})$ $b_2 = \log(b_1)$
P8	$R = E \exp(a_4 - \exp(b_2)(E - \bar{E}))$	$a_4 = \log(a_1) - b_1 \bar{E}$ $b_2 = \log(b_1)$
P9	$R = a_1 E (b_3)^{-E/1000}$	$b_3 = \exp(1000 b_1)$
P10	$R = E \exp(a_2)(b_3)^{-E/1000}$	$a_2 = \log(a_1)$ $b_3 = \exp(1000 b_1)$
P11	$R = a_3 E (b_3)^{-(E-\bar{E})/1000}$	$a_3 = a_1 \exp(- b_1 \bar{E})$ $b_3 = \exp(1000 b_1)$
P12	$R = E \exp(a_4)(b_3)^{-(E-\bar{E})/1000}$	$a_4 = \log(a_1) - b_1 \bar{E}$ $b_3 = \exp(1000 b_1)$

parameterisation, data set combination are given in Table 4.5. The simulation and asymmetry measures of nonlinearity for the seven parameters a_1, a_2, a_3, a_4, b_1, b_2, b_3 are given in Table 4.6.

All the parameterisations, except for P5, have lower parameter effects curvature than the original parameterisation P1 for all three data sets. (It is interesting to note that P5 incorporates one of the recommended transformations (ie transformation iv above), showing that "good" parameterisations are specific to model, data set combinations). In particular, the parameter effects curvatures of P2, P3 and P4 are consistently low compared to those of P1. The simulation and asymmetry measures in Table 4.6 show that none of the estimators of the parameters of P2, P3 and P4 have large nonlinear behaviour. Since P2 is the only one of the three parameterisations whose parameters, namely a_2 and b_1, are measures of density independent and density dependent survival, it is concluded that P2,

$$R_t = E_t \exp(\alpha_2 - b_1 E_t) + \epsilon_t, \qquad (4.5)$$

is the best parameterisation of the additive error Ricker model for the R_1, R_1' and R_2' data sets.

The value of using parameterisation (4.5) is demonstrated by comparing an exact and a conservative linear approximation 95% confidence region for a_2 and b_1 for the R_1' data set (Fig. 4.4); the linear approximation in a_2 and b_1 is very good. Further, a linear approximation 95% confidence region for a_1 and b_1 is obtained by mapping the linear approximation confidence region for a_2 and b_1 into (a_1,b_1)

TABLE 4.5

Parameter effects curvature

R_1 (May/June)					
		a_1	a_2	a_3	a_4
b_1	max	0.068	0.039	0.034	0.039
	RMS	0.044	0.027	0.022	0.027
b_2	max	0.116	0.037	0.037	0.037
	RMS	0.070	0.025	0.023	0.031
b_3	max	0.052	0.058	0.044	0.045
	RMS	0.039	0.040	0.027	0.030
R_1' (May/June)					
		a_1	a_2	a_3	a_4
b_1	max	0.199	0.113	0.100	0.113
	RMS	0.129	0.079	0.066	0.079
b_2	max	0.340	0.109	0.107	0.108
	RMS	0.205	0.074	0.068	0.089
b_3	max	0.153	0.169	0.129	0.133
	RMS	0.115	0.116	0.079	0.087
R_2' (Aug/Sep)					
		a_1	a_2	a_3	a_4
b_1	max	0.146	0.058	0.052	0.058
	RMS	0.089	0.045	0.035	0.045
b_2	max	0.267	0.107	0.073	0.076
	RMS	0.160	0.067	0.044	0.053
b_3	max	0.119	0.086	0.063	0.068
	RMS	0.076	0.064	0.040	0.049

TABLE 4.6

Simulation and asymmetry measures of nonlinearity

	R_1	R_1'	R_2'
Parameter: a_1			
Least squares estimate	$4.775*10^{-1}$	$4.939*10^{-1}$	$8.189*10^{-2}$
Lin approx s.e.	$0.133*10^{-1}$	$0.404*10^{-1}$	$0.467*10^{-2}$
Box's bias	$0.002*10^{-1}$	$0.019*10^{-1}$	$0.011*10^{-2}$
Hougaard's skewness	0.075	0.219 *	0.110
Lowry and Morton's λ	0.001	0.006	0.001
Simulation measures:			
True value	$4.775*10^{-1}$	$4.775*10^{-1}$	$8.189*10^{-2}$
Variance of l.s.e (lin approx)	$1.774*10^{-4}$	$1.277*10^{-3}$	$2.183*10^{-5}$
Simulated value	$4.781*10^{-1}$	$4.880*10^{-1}$ *	$8.2000*10^{-2}$
Simulated variance	$1.762*10^{-4}$	$1.420*10^{-3}$ *	$2.2660*10^{-5}$
T_3 - skewness stat	1.56	1.91	2.15 *
T_4 - excess kurtosis stat	0.26	2.85 **	2.13 *
Lowry and Morton's λ	0.001	0.005	0.002
	R_1	R_1'	R_2'
Parameter: a_2			
Least squares estimate	$-7.391*10^{-1}$	$-7.054*10^{-1}$	-2.502
Lin approx s.e.	$0.279*10^{-1}$	$0.818*10^{-1}$	0.057
Box's bias	$0.001*10^{-1}$	$0.004*10^{-1}$	$-3*10^{-4}$
Hougaard's skewness	-0.009	-0.026	-0.061
Lowry and Morton's λ	$4*10^{-4}$	0.004	0.001
Simulation measures:			
True value	$-7.391*10^{-1}$	$-7.391*10^{-1}$	-2.502
Variance of l.s.e (lin approx)	$7.779*10^{-4}$	$5.600*10^{-3}$	$3.255*10^{-3}$
Simulated value	$-7.382*10^{-1}$	$-7.371*10^{-1}$	-2.503
Simulated variance	$7.698*10^{-4}$	$6.204*10^{-3}$ *	$3.369*10^{-3}$
T_3 - skewness stat	0.47	-1.88	-0.40
T_4 - excess kurtosis stat	0.10	3.80 **	1.62
Lowry and Morton's λ	0.001	0.003	0.001

TABLE 4.6 (continued)

	R_1	R_1'	R_2'
Parameter: a_3			
Least squares estimate	$9.330*10^{-2}$	$9.761*10^{-2}$	$2.661*10^{-2}$
Lin approx s.e.	$0.141*10^{-2}$	$0.431*10^{-2}$	$0.068*10^{-2}$
Box's bias	$-0.001*10^{-2}$	$-0.010*10^{-2}$	$-0.001*10^{-2}$
Hougaard's skewness	-0.013	-0.038	$-1*10^{-5}$
Lowry and Morton's λ	$2*10^{-4}$	0.002	$3*10^{-4}$
Simulation measures:			
True value	$9.330*10^{-2}$	$9.330*10^{-2}$	$2.661*10^{-2}$
Variance of l.s.e (lin approx)	$1.991*10^{-6}$	$1.433*10^{-5}$	$4.575*10^{-6}$
Simulated value	$9.332*10^{-2}$	$9.303*10^{-2}$ *	$2.658*10^{-2}$
Simulated variance	$2.123*10^{-6}$	$1.455*10^{-5}$	$4.815*10^{-6}$
T_3 - skewness stat	1.06	-0.03	-0.45
T_4 - excess kurtosis stat	0.42	0.10	-0.67
Lowry and Morton's λ	$3*10^{-4}$	0.002	0.001

	R_1	R_1'	R_2'
Parameter: a_4			
Least squares estimate	-2.372	-2.326	-3.626
Lin approx s.e.	0.015	0.044	0.025
Box's bias	$-2*10^{-4}$	-0.002	$-6*10^{-4}$
Hougaard's skewness	-0.059	-0.171	-0.076
Lowry and Morton's λ	$4*10^{-4}$	0.003	0.001
Simulation measures:			
True value	-2.372	-2.372	-3.626
Variance of l.s.e (lin approx)	$2.287*10^{-4}$	$1.647*10^{-3}$	$6.460*10^{-4}$
Simulated value	-2.372	-2.376 **	-3.628
Simulated variance	$2.434*10^{-4}$	$1.689*10^{-3}$	$6.833*10^{-4}$
T_3 - skewness stat	0.44	-1.63	-1.42
T_4 - excess kurtosis stat	0.33	0.19	-0.52
Lowry and Morton's λ	0.001	0.003	0.001

TABLE 4.6 (continued)

	R_1	R_1'	R_2'
Parameter: b_1			
Least squares estimate	$4.065*10^{-4}$	$4.036*10^{-4}$	$2.658*10^{-4}$
Lin approx s.e.	$0.082*10^{-4}$	$0.238*10^{-4}$	$0.122*10^{-4}$
Box's bias	$0.001*10^{-4}$	$0.006*10^{-4}$	$0.001*10^{-4}$
Hougaard's skewness	0.052	0.153	0.037
Lowry and Morton's λ	0.001	0.005	0.001
Simulation measures:			
True value	$4.065*10^{-4}$	$4.065*10^{-4}$	$2.658*10^{-4}$
Variance of l.s.e (lin approx)	$6.674*10^{-11}$	$4.805*10^{-10}$	$1.478*10^{-10}$
Simulated value	$4.067*10^{-4}$	$4.079*10^{-4}$ *	$2.661*10^{-4}$
Simulated variance	$6.517*10^{-11}$	$5.222*10^{-10}$	$1.498*10^{-10}$
T_3 - skewness stat	0.91	1.68	-0.20
T_4 - excess kurtosis stat	0.11	0.62	0.11
Lowry and Morton's λ	0.001	0.004	0.002

	R_1	R_1'	R_2'
Parameter: b_2			
Least squares estimate	-7.808	-7.815	-8.233
Lin approx s.e.	0.020	0.059	0.046
Box's bias	$-3*10^{-5}$	$-2*10^{-4}$	-0.001
Hougaard's skewness	-0.008	-0.024	-0.010
Lowry and Morton's λ	0.001	0.004	0.002
Simulation measures:			
True value	-7.808	-7.808	-8.233
Variance of l.s.e (lin approx)	$4.039*10^{-4}$	$2.908*10^{-3}$	$2.091*10^{-4}$
Simulated value	-7.808	-7.806	-8.233
Simulated variance	$3.939*10^{-4}$	$3.141*10^{-3}$	$2.219*10^{-4}$
T_3 - skewness stat	0.14	-0.58	-2.02 *
T_4 - excess kurtosis stat	0.06	0.58	0.46
Lowry and Morton's λ	0.001	0.004	0.002

TABLE 4.6 (continued)

	R_1	R_1'	R_2'
Parameter: b_3			
Least squares estimate	1.502	1.497	1.304
Lin approx s.e.	0.012	0.036	0.016
Box's bias	$1*10^{-4}$	0.001	$2*10^{-4}$
Hougaard's skewness	0.077	0.224 *	0.074
Lowry and Morton's λ	0.001	0.007	0.002
Simulation measures:			
True value	1.502	1.502	1.304
Variance of l.s.e (lin approx)	$1.505*10^{-4}$	$1.083*10^{-3}$	$2.514*10^{-4}$
Simulated value	1.502	1.504 *	1.305
Simulated variance	$1.471*10^{-4}$	$1.185*10^{-3}$ *	$2.551*10^{-4}$
T_3 - skewness stat	1.23	2.61 **	0.27
T_4 - excess kurtosis stat	0.16	0.84	0.10
Lowry and Morton's λ	0.001	0.006	0.002

* indicates moderate nonlinear behaviour
** indicates serious nonlinear behaviour

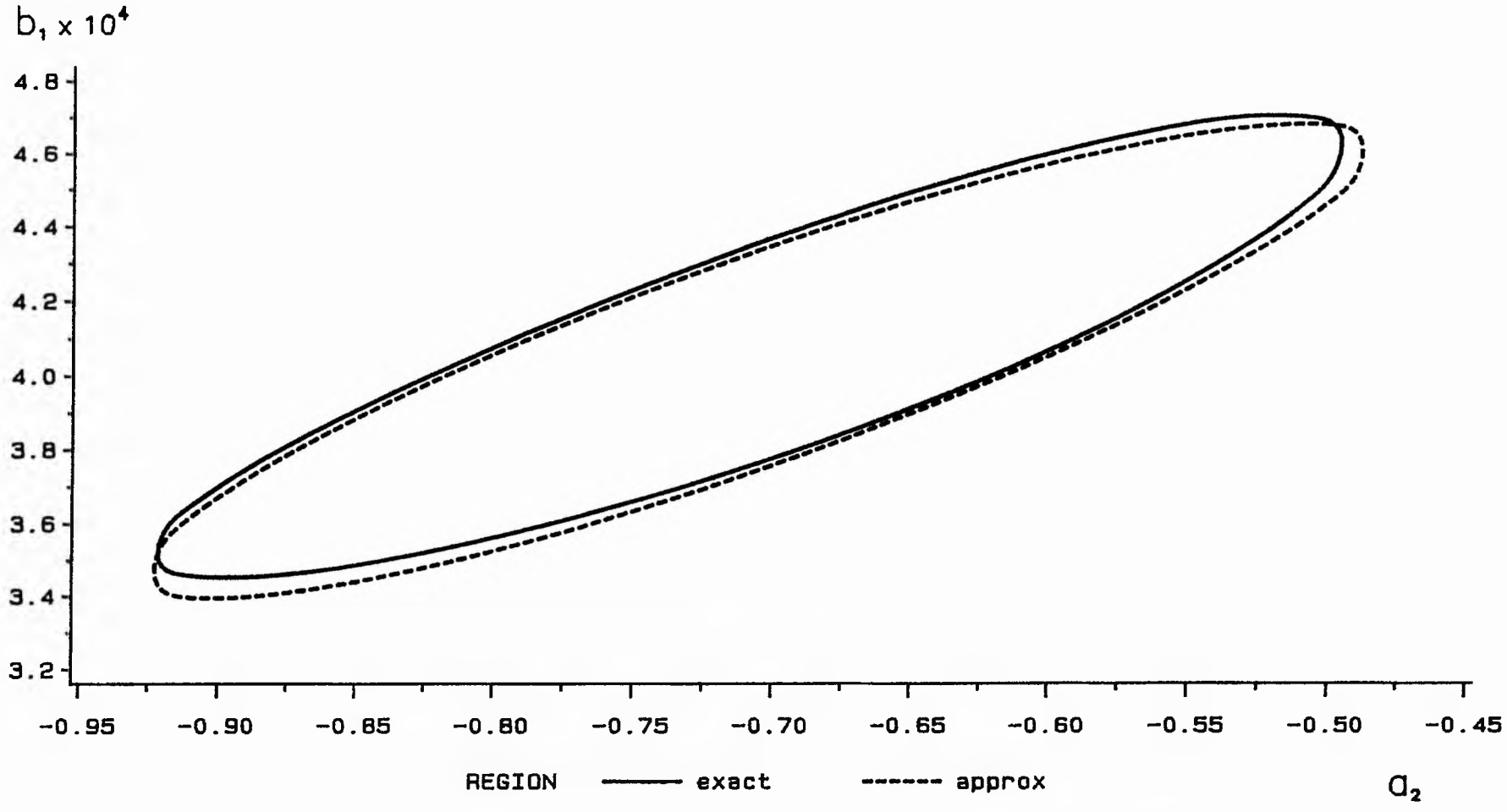


Fig. 4.4 Exact and linear approximation 95% confidence regions for a_2 and b_1 (R_1' data set).

space (Fig. 4.5); this gives a better approximation than that obtained from the original parameterisation (4.4) (Fig. 4.3).

4.4 ZERO PARAMETER EFFECTS CURVATURE

The additive error Ricker model is a two parameter model which can be parameterised so that it is conditionally linear in one of its parameters. That is, when written as

$$R_i = a_1 E_i \exp(-b_1 E_i) + \epsilon_i,$$

the Ricker model is conditionally linear in a_1 (since $\partial R/\partial a_1$ does not depend on a_1). Hence, there exists a parameterisation of the Ricker model for which the parameter effects curvature is zero (Bates and Watts, 1981).

This parameterisation is obtained by a method described in Bates and Watts (1981). The new parameters γ_1 and γ_2 are given by

$$\gamma_1 = a_1 u \exp(-b_1 u)$$

$$\gamma_2 = a_1 v \exp(-b_1 v)$$

and are interpreted as the expected number of recruits obtained from u and v eggs respectively. The values of u and v are found by equating the arrays

$$T = \begin{pmatrix} 0 & 0 \\ 0 & -a_1 uv \end{pmatrix} \begin{pmatrix} 0 & 1/a_1 \\ 1/a_1 & -(u+v) \end{pmatrix}$$

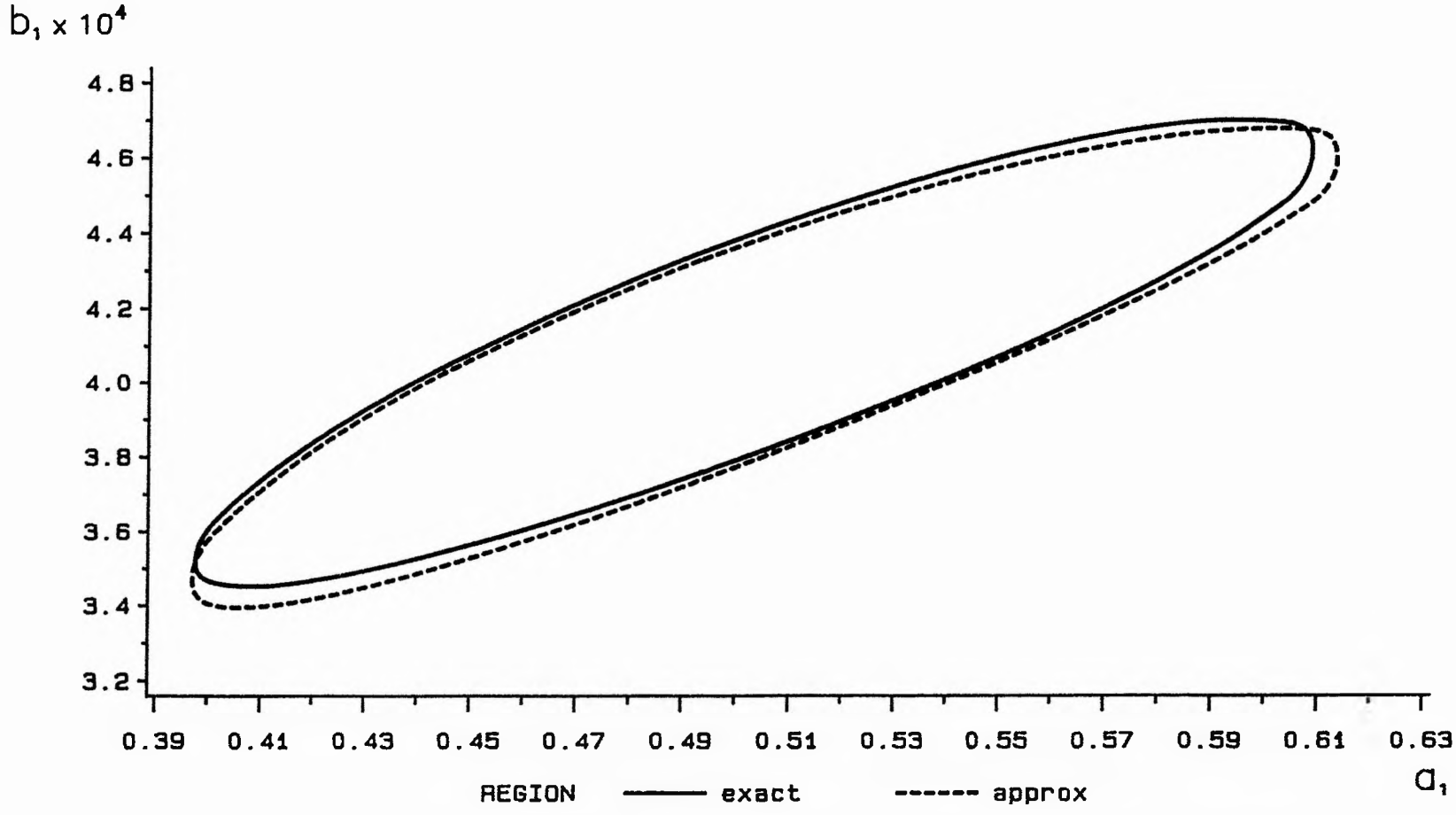


Fig. 4.5 An exact 95% confidence region for a_1 and b_1 and an approximate 95% confidence region for a_1 and b_1 obtained by mapping a linear approximation 95% confidence region for a_2 and b_1 into (a_1, b_1) space (R_1' data set).

and

$$T^*(\hat{a}_1,\hat{b}_1) = \begin{pmatrix} 0 & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} 0 & 1/\hat{a}_1 \\ 1/\hat{a}_1 & Q \end{pmatrix}$$

where, in Bates and Watts' notation,

i) $\hat{a}_1$ and $\hat{b}_1$ are the least squares estimates of a_1 and b_1,

ii) T is the transformation array,

iii) $T^*(\hat{a}_1,\hat{b}_1)$ is the target transformation array,

iv) $P = ([(V.'V.)^{-1} V.'] [V..])_{122}$

$Q = ([(V.'V.)^{-1} V.'] [V..])_{222}$

where $V.$ and $V..$ are the arrays of (scaled) first and second partial derivatives respectively.

Thus

$$u = (-Q - \sqrt{(Q^2 + 4P/\hat{a}_1)})/2,$$

$$v = (-Q + \sqrt{(Q^2 + 4P/\hat{a}_1)})/2.$$

The values of u and v for the three data sets are given in Table 4.7.

Unfortunately, this parameterisation has a number of limitations. First, the Ricker model, expressed in terms of γ_1 and γ_2

TABLE 4.7

Values of u and v

Data set	u	v
R_1	2205	6002
R_1'	2211	6024
R_2'	2564	6870

$$R_i = E_i(\gamma_1/u)^{(v-E_i)/(v-u)} \quad (\gamma_2/v)^{(E_i-u)/(v-u)} + \epsilon_i$$

gives no intuitive feel for the relationship between R and E. Secondly, the parameters γ_1 and γ_2 can not be interpreted as measures of density independent and density dependent survival. Thirdly, there is no simple way of adapting the parameterisation when the Ricker model is extended to include other covariates, such as the measure of drought used in Chapter 6. Therefore, this parameterisation is not considered further.

4.5 DISCUSSION

In this Chapter, it is shown that (4.5) is the best parameterisation of the additive error Ricker model for the R_1, R_1' and R_2' data sets. However, it is purely coincidental that this is the same parameterisation as that of the "linearised" form of the multiplicative error Ricker model (4.3). In general, it is not necessarily true that if the nonlinear model

$$y_i = f(x_i, \theta) + \epsilon_i, \quad i = 1 \ldots n$$

can be written as a function of a linear combination of its parameters

$$y_i = f(\sum_{j=1}^{p} x_{ij}\theta_j) + \epsilon_i, \quad i = 1 \ldots n, \tag{4.6}$$

then the nonlinear behaviour of the model will be small (Ratkowsky, 1983). For example, consider model (3.7a)

$$y_i = \frac{x_i}{1 + \theta x_i} + \epsilon_i, \quad i = 1,2$$

This is of the same form as (4.6) since it can be written as

$$y_i = (\theta + 1/x_i)^{-1} + \epsilon_i, \quad i = 1,2.$$

However, the nonlinear behaviour of the model with this parameterisation is large. Indeed, the reparameterisation (3.7b)

$$y_i = \frac{\phi x_i}{\phi + x_i} + \epsilon_i, \quad i = 1,2 \tag{4.7}$$

considerably reduces the nonlinear behaviour, although (4.7) can not be written in the same form as (4.6).

Since the additive and the multiplicative error Ricker models have the same parameterisation, it is straightforward to compare inferences from each model. For example, 95% confidence regions for a_2 and b_1 (R_1 data set) obtained from each model are shown in Figure 4.6. These regions are fairly similar and suggest that the choice of error structure does not have a large effect on inferences about the parameters a_2 and b_1.

The additive error Ricker model is parameterised as (4.5) throughout the rest of this Thesis. However, since the model (or extensions of the model) are applied to all the life stages R_1, R_2, R_3 and R_4 and since it is important to distinguish between the parameter values for each life stage, it is convenient to adopt a change of

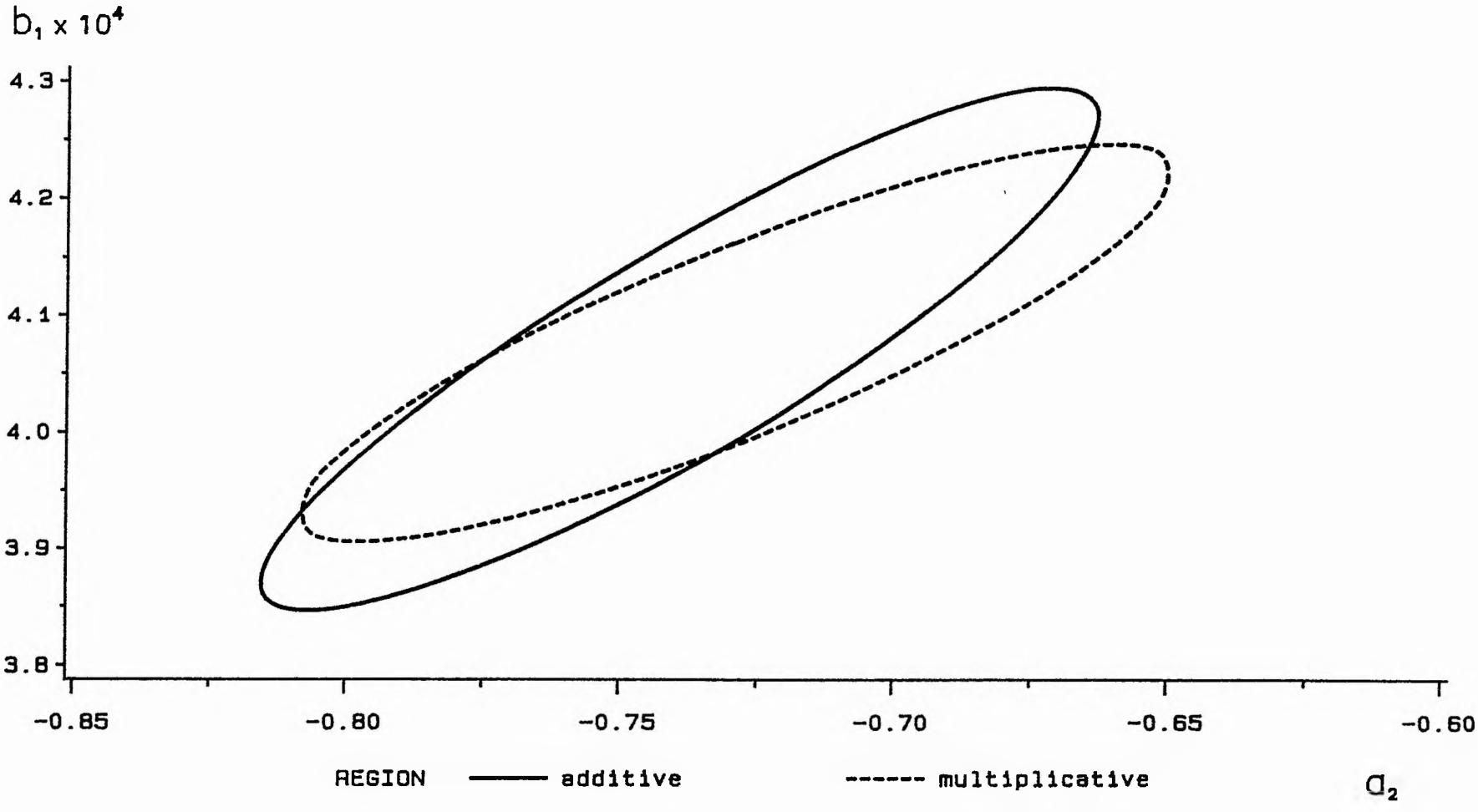


Fig. 4.6 Exact 95% confidence regions for a_2 and b_1 for both the additive and the multiplicative error Ricker models (R_1 data set).

notation. Let α_j and β_j be the parameters a_2 and b_1 respectively, when the Ricker model is applied to the life stage R_j, $j = 1...4$. Thus, the Ricker model applied to the R_1 data set is written

$$R_{1i} = E_i \exp(\alpha_1 - \beta_1 E_i) + \epsilon_i, \quad i = 67...84; \tag{4.8}$$

the model applied to the R_2 data set is written

$$R_{2i} = E_i \exp(\alpha_2 - \beta_2 E_i) + \epsilon_i, \quad i = 67...84; \tag{4.9}$$

and so on.

CHAPTER 5 ERRORS-IN-VARIABLES AND LEAST SQUARES ESTIMATION

5.1 INTRODUCTION

Most standard regression theory assumes that the independent variates are known without error. For example, in Chapter 4, the distributions of the least squares estimators and the confidence regions for (α_1,β_1) are derived on the assumption that there are no errors associated with the values $\{E_i\}$. This assumption is not strictly valid, since the number of eggs is only estimated from the number of redds cut in a given year. Hence, for example, the least squares estimators will not necessarily be unbiased (approximately so in the nonlinear case) nor have a variance close to the minimum variance bound. This Chapter investigates the effect of the errors in the egg variable on the distributions of the least squares estimators of (α_1,β_1).

An introduction to the theory of errors-in-variables and the ways such errors affect least squares estimation is given by Draper and Smith (1981). Kendall and Stuart (1979, Chapter 29) give a much more detailed discussion of the general errors-in-variables problem. Nonlinear errors-in-variables (such as occur in the Ricker model) are more complex; Wolter and Fuller (1982) consider the quadratic errors-in-variables problem and give references to work in the field. In fisheries research, Uhler (1980), Walters and Ludwig (1981), Ludwig and Walters (1981) and Walters (1985) consider errors-in-variables in various production and stock-recruitment models (see Chapter 2). This chapter describes how errors-in-variables arise in the Black Brows Beck data and how they affect the least squares estimation of (α_1,β_1). In Section 5.2, the method of estimating the number

of eggs is described in detail and the errors-in-variables Ricker model is formulated. Section 5.3 derives an approximate formula for $E(\tilde{\beta}_1)$, the expectation of the least squares estimator of β_1 in the multiplicative error Ricker model (4.3). This formula, and similar formulae for the expectations and variances of the other least squares estimators, is very complicated. Hence, in Section 5.4, simplifying assumptions are made which enable approximate distributions of the least squares estimators to be derived. These distributions are also investigated by a series of simulations.

5.2 THE ERRORS-IN-VARIABLES RICKER MODEL

5.2.1 Errors in the Egg Variable

The true number of eggs E_i in year i is estimated as follows. The spawning season is divided into three periods - before 14 November, 14 to the 25 November and after 25 November - which roughly correspond to the spawning times of large, medium and small trout respectively. The number of eggs in a redd cut in period j is assumed to be distributed $N(\mu_j, \tau_j^2)$, $j = 1...3$. Between 1967 and 1981, n_j redds cut in period j in the lower section of the stream were excavated and the eggs in each redd were counted (Elliott, 1984a); μ_j and τ_j^2 are estimated by the sample mean $\hat{\mu}_j$ and variance $\hat{\tau}_j^2$ of these counts, $j = 1...3$ (Table 5.1). No evidence was found to discount the assumption of normality.

Let z_{ij} be the number of redds cut in period j in year i, $j = 1...3$, $i = 67...84$. The numbersof redds are assumed to be counted without error. Then E_i is estimated by e_i, where

$$e_i = \sum_{j=1}^{3} z_{ij}\hat{\mu}_j, \qquad i = 67...84.$$

TABLE 5.1

Estimates of the numbers of eggs per redd

Time period	Eggs per redd		
	$\hat{\mu}$	$\hat{\tau}^2$	n
- 13 November	1550	$2.765*10^4$	10
14 November - 25 November	1034	$2.163*10^4$	16
26 November -	722	$1.337*10^4$	17

The values of (z_{ij}) and (e_i) are given in Table 5.2.

Two types of errors are involved in the estimation of (E_i); these arise as follows. Let q_i be the estimated number of eggs in year i if (μ_j) are known,

$$q_i = \sum_{j=1}^{3} z_{ij}\mu_j, \quad i = 67...84,$$

and let

$$\delta_i = E_i - q_i, \quad i = 67...84. \tag{5.1}$$

Also, let

$$\hat{\mu}_j = \mu_j + \xi_j, \quad j = 1...3,$$

so that

$$e_i = \sum_{j=1}^{3} z_{ij}\hat{\mu}_j = \sum_{j=1}^{3} z_{ij}\mu_j + \sum_{j=1}^{3} z_{ij}\xi_j$$

$$= q_i + \nu_i, \quad i = 67...84, \tag{5.2}$$

where

$$\nu_i = \sum_{j=1}^{3} z_{ij}\xi_j, \quad i = 67...84. \tag{5.3}$$

TABLE 5.2

Estimates of error variances and the number of eggs

Year	z_1	z_2	z_3	e	Var(υ/q) $*10^3$	Var(δ/q) $*10^2$	Var(δ) $*10^{-4}$	Var(υ) $*10^{-4}$	Var($\delta\beta_{10}$) $*10^2$
1967	1	3	0	4652	0.69	0.43	9.38	1.51	1.55
1968	1	0	1	2272	0.69	0.79	4.14	0.36	0.68
1969	0	1	1	1756	0.69	1.14	3.52	0.22	0.58
1970	1	1	0	2584	0.62	0.74	5.00	0.42	0.83
1971	0	1	5	4644	0.97	0.41	8.85	2.10	1.46
1972	0	0	4	2888	1.51	0.64	5.33	1.25	0.88
1973	0	0	5	3610	1.51	0.51	6.66	1.96	1.10
1974	2	4	1	7958	0.52	0.25	15.71	3.39	2.60
1975	2	2	3	7334	0.44	0.26	14.00	2.38	2.31
1976	0	2	3	4234	0.70	0.47	8.38	1.25	1.38
1977	2	1	3	6300	0.49	0.29	11.81	1.97	1.95
1978	0	2	0	2068	1.26	1.01	4.38	0.55	0.72
1979	0	1	0	1034	1.26	2.02	2.19	0.14	0.36
1980	0	0	1	722	1.51	2.57	1.33	0.08	0.22
1981	2	3	2	7646	0.45	0.25	14.85	2.67	2.45
1982	2	4	1	7958	0.53	0.25	15.71	3.39	2.60
1983	0	1	2	2478	0.73	0.79	4.86	0.45	0.80
1984	0	0	3	2166	1.51	0.85	4.00	0.71	0.66

Thus, from (5.1) and (5.2),

$$e_i = E_i - \delta_i + \nu_i, \quad i = 67...84. \tag{5.4}$$

The errors $\{\nu_i\}$ are due to the sampling errors in estimating $\{\mu_j\}$ from the samples of redds cut lower down the stream; these errors can be reduced by sampling more redds (ie by increasing $\{n_j\}$) (see Section 5.2.2). The errors $\{\delta_i\}$ are due to the variation of the numbers of eggs in the redds in the study section about the values $\{\mu_j\}$; these errors are inherent to the method of estimation and can not be reduced.

It is possible that there are also non-sampling errors due to the use of data from the lower section of the stream to estimate the numbers of eggs per redd in the upper study section. However, it is assumed that these errors are negligible.

5.2.2 The Distributions of the Errors

It is assumed that the number of recruits R_1 is related to the true number of eggs E (rather than the estimated number of eggs e) by a Ricker function with either an additive error

$$R_{1i} = E_i \exp(\alpha_1 - \beta_1 E_i) + \epsilon_i, \quad i = 67...84, \tag{5.5a}$$

or a multiplicative error

$$\log(R_{1i}) = \log(E_i) + \alpha_1 - \beta_1 E_i + \epsilon_i, \quad i = 67...84, \tag{5.5b}$$

where

$$\{\epsilon_i\} \text{ are } NID(0,\sigma^2). \tag{5.6}$$

(It is assumed that movement of the young fish into and out of the study section is negligible, so that the recruits R_1 all originate from the eggs laid in the study section, rather than from eggs laid elsewhere in the stream.) Substituting (5.4) into (5.5) gives

$$R_{1i} = (e_i + \delta_i - \nu_i)\exp(\alpha_1 - \beta_1(e_i + \delta_i - \nu_i)) + \epsilon_i, \quad i = 67...84, \tag{5.7a}$$

$$\log(R_{1i}) = \log(e_i + \delta_i - \nu_i) + \alpha_1 - \beta_1(e_i + \delta_i - \nu_i) + \epsilon_i, \quad i = 67...84. \tag{5.7b}$$

The least squares estimates $(\hat{\alpha}_1,\hat{\beta}_1)'$ and $(\tilde{\alpha}_1,\tilde{\beta}_1)'$ are obtained from (5.7a) and (5.7b) respectively, by assuming that $\delta_i = 0$ and $\nu_i = 0$, $i = 67...84$. However, the distributions of the least squares estimators depend on the distributions of all three types of "errors" $\{\epsilon_i\}$, $\{\nu_i\}$ and $\{\delta_i\}$, where

i) the distribution of $\{\epsilon_i\}$ is given by (5.6),

ii) the distribution of $\{\nu_i\}$ depends on the distribution of $\{\xi_j\}$ through (5.3). The number of eggs in a redd cut in period j is distributed $N(\mu_j,\tau_j^2)$. It is assumed that the numbers of eggs in different redds are independent, so that $\{\xi_j\}$ are independent with

$$\xi_j \sim N(0, \tau_j^2/n_j), \quad j = 1...3.$$

Thus,

$$v_i \sim N(0, \sum_{j=1}^{3} z_{ij}^2(\tau_j^2/n_j)), \quad i = 67...84,$$

$$(\text{with } \mathrm{Cov}(v_i, v_k) = \sum_{j=1}^{3} z_{ij} z_{kj}(\tau_j^2/n_j), \quad i = 67...84, \; k = 67...84),$$

iii) strictly, $\{E_i\}$ and $\{q_i\}$ and hence $\{\delta_i\}$ are constant since, for example, E_{67} is the true number of eggs laid in Black Brows Beck in 1967. Thus, the distributions of $(\hat{\alpha}_1, \hat{\beta}_1)'$ and $(\tilde{\alpha}_1, \tilde{\beta}_1)'$ are functions of the values of $\{E_i\}$ and $\{q_i\}$ (or $\{\delta_i\}$). However, these functions are of limited use since the values $\{E_i\}$ are unknown. An alternative approach is to take a superpopulation view of sampling. The true numbers of eggs $\{E_i\}$ are regarded as one realisation of a multivariate normal distribution with mean

$$q = (q_{67}, \ldots, q_{84})'$$

and covariance matrix

$$\mathrm{diag}(\sum_{j=1}^{3} z_{67,j}\tau_j^2, \ldots, \sum_{j=1}^{3} z_{84,j}\tau_j^2).$$

Then, $\{\delta_i\}$ are independent, with

$$\delta_i \sim N(0, \sum_{j=1}^{3} z_{ij}\tau_j^2), \quad i = 67...84. \qquad (5.8)$$

It is possible that there are correlations between γ_i, δ_i and ε_i within years. This might occur if, for example, there are strong year class effects. However, there is no way of assessing whether such correlations exist, so it is assumed that the three sets of errors $\{\gamma_i\}$, $\{\delta_i\}$ and $\{\varepsilon_i\}$ are independent. (This is a reasonable assumption if year class effects can be ignored, since $\{\xi_j\}$ and hence $\{\gamma_i\}$ are errors due to the variability of the number of eggs in redds in the lower study section, $\{\delta_i\}$ are errors due to the variability of the number of eggs in redds in the upper study section and $\{\varepsilon_i\}$ are errors due to the variability of the number of recruits).

5.3 THE EXPECTATION OF $\tilde{\beta}_1$

In this section, an approximate formula for $E(\tilde{\beta}_1)$ is derived. To simplify the presentation, some distributions that are required of functions of the errors $\{\gamma_i\}$ are derived in Appendix 5 at the end of this Chapter.

The least squares estimator of β_1 in the multiplicative error model (5.7b) is given by

$$\tilde{\beta}_1 = -\frac{\sum_{i=67}^{84}\{(\log(R_{1i})-\log(e_i)-\frac{1}{18}\sum_{k=67}^{84}\{\log(R_{1k})-\log(e_k)\})(e_i-\frac{1}{18}\sum_{k=67}^{84}e_k)\}}{\sum_{i=67}^{84}(e_i-\frac{1}{18}\sum_{k=67}^{84}e_k)^2}$$

$$= -S[\log(R_1)-\log(e),e]/S[e,e] \qquad (5.9)$$

where, if u and v are the n-vectors $u = (u_1,\ldots,u_n)'$ and $v = (v_1,\ldots,v_n)'$, then

$$\bar{u} = \frac{1}{n}\sum_{i=1}^{n}u_i, \quad S[u,v] = \sum_{i=1}^{n}(u_i - \bar{u})(v_i - \bar{v}).$$

It is convenient to express (5.9) in terms of the errors (ε_i) and (ν_i) and the values (q_i). First, the denominator is expanded. Since $e_i = q_i + \nu_i$, $i = 67...84$,

$$S[e,e] = \sum_{i=67}^{84}(q_i + \nu_i - \bar{q} - \bar{\nu})^2 = S[q,q] + 2S[q,\nu] + S[\nu,\nu]$$

$$= S[q,q]\left(1 + \frac{2S[q,\nu]}{S[q,q]} + \frac{S[\nu,\nu]}{S[q,q]}\right)$$

From (A5.4) and (A5.5) (see Appendix 5)

$$E(S[q,\nu]) = 0$$

$$Var(S[q,\nu]) = \sum_{j=1}^{3}(\tau_j^2/n_j)S[q,z_j]^2,$$

where $z_j = (z_{67,j},...,z_{84,j})'$, $j = 1...3$. Approximating (μ_j) and (τ_j^2) by their sample estimates $(\hat{\mu}_j)$ and $(\hat{\tau}_j^2)$ gives

$$Var(S[q,\nu]/S[q,q]) \approx 5.08*10^{-4}$$

Thus, an approximate 95% interval for $S[q,\nu]/S[q,q]$ is (-0.044, 0.044), so that $S[q,\nu]/S[q,q]$ is small compared to 1 with high probability. Similarly, from (A5.6) and (A5.9),

$$E(S[\nu,\nu]) = \sum_{j=1}^{3}(\tau_j^2/n_j)S[z_j,z_j]$$

$$\mathrm{Var}(S[v,v]) = \sum_{j=1}^{3}\sum_{k=1}^{3}(\tau_j^2/n_j)(\tau_k^2/n_k)\{S[z_j,z_j]S[z_k,z_k] + 2S[z_j,z_k]^2\} - \{E(S[v,v])\}^2$$

Thus,

$$E(S[v,v]/S[q,q]) \approx 1.13*10^{-3},$$

$$\mathrm{Var}(S[v,v]/S[q,q]) \approx 1.19*10^{-6}$$

so that $S[v,v]/S[q,q]$ is small compared to 1 with high probability. Hence, the term $1/S[e,e]$ in equation (5.9) can be approximated by

$$\frac{1}{S[e,e]} \approx \frac{1}{S[q,q]}\left(1 - \frac{2S[q,v]}{S[q,q]} - \frac{S[v,v]}{S[q,q]}\right) \qquad (5.10)$$

The numerator in (5.9) is rewritten as follows. First,

$$\log(e_i) = \log(q_i + v_i) = \log(q_i) + \log(1 + v_i/q_i), \quad i = 67...84.$$

Now,

$$E(v_i/q_i) = E(\sum_{j=1}^{3} z_{ij}\xi_j/q_i) = 0, \quad i = 67...84,$$

$$\mathrm{Var}(v_i/q_i) = \sum_{j=1}^{3}(\tau_j^2/n_j)(z_{ij}/q_i)^2 < 1.51*10^{-3}, \quad i = 67...84,$$

(see Table 5.2), so ν_i/q_i is small compared to 1 with high probability, $i = 67...84$. Thus,

$$\log(e_i) \cong \log(q_i) + \nu_i/q_i, \quad i = 67...84,$$

and, using (5.5b),

$$S[\log(R_1) - \log(e), e]$$

$$\cong S[\log(E) + \alpha_1 - \beta_1 E + \epsilon - \log(q) - \nu/q, q + \nu].$$

Since α_1 is a constant,

$$S[\log(R_1) - \log(e), e] \cong S[\log(E) - \beta_1 E + \epsilon - \log(q) - \nu/q, q + \nu]. \tag{5.11}$$

Combining (5.9), (5.10) and (5.11) gives

$$\beta_1 \cong (A + B + C + D)/S[q, q], \tag{5.12}$$

where

$$A = S[\beta_1 E + \log(q/E), q],$$

$$B = -S[\epsilon, q + \nu]\{1 - (2S[q, \nu] + S[\nu, \nu])/S[q, q])\},$$

$$C = S[\beta_1 E + \log(q/E), \nu] + S[\nu/q, q] -$$

$$2S[\beta_1 E + \log(q/E), q]S[q, \nu]/S[q, q],$$

$$S[q,q]D = S[q,q]S[\nu/q,\nu] - S[\beta_1 E + \log(q/E),\ q]S[\nu,\nu] -$$

$$\{S[\beta_1 E + \log(q/E),\ \nu] + S[\nu/q, q + \nu]\}\{S[q,\nu] + S\{\nu,\nu]\}.$$

The term A is independent of the errors $\{\epsilon_i\}$ and $\{\nu_i\}$, B is a linear combination of $\{\epsilon_i\}$, C is a linear combination of $\{\nu_i\}$ and D contains the remainder of the expression. The expectation of $\tilde{\beta}_1$ can now be expressed fairly simply. First, $E(A) = A$, since A is independent of $\{\epsilon_i\}$ and $\{\nu_i\}$. Also $E(B) = 0$ and $E(C) = 0$ since $\{\epsilon_i\}$ are independent of $\{\nu_i\}$ and $E(\epsilon_i) = 0$, $E(\nu_i) = 0$, $i = 67...84$. Thus

$$E(\tilde{\beta}_1) = \{A + E(D)\}/S[q,q],$$

where, using (A5.6), (A5.7), (A5.8) and (A5.9),

$$S[q,q]E(D) = \sum_{j=1}^{3}(\tau_j^2/n_j)\{S[q,q]S[z_j/q,z_j] -$$

$$S[\beta_1 E + \log(q/E),\ q]S[z_j,z_j] -$$

$$2S[q,z_j](S[\beta_1 E + \log(q/E),\ z_j] - S[z_j/q,q]) +$$

$$\sum_{k=1}^{3}(\tau_k^2/n_k)(S[z_j/q,z_j]S[z_k,z_k] + 2S[z_j/q,z_k]S[z_j,z_k])\}$$

However, if the values of $\{E_i\}$ are unknown (as is the case here), it is not possible to estimate either $E(\tilde{\beta}_1)$ or the bias of $\tilde{\beta}_1$.

If $\{E_i\}$ are regarded as a realisation of a multivariate normal distribution, then (5.12) can be simplified and an estimate of the bias of $\tilde{\beta}_1$ can be obtained. In this case

$$E_i = q_i + \delta_i, \quad i = 67...84,$$

where $\{\delta_i\}$ are independent with $\delta_i \sim N(0, \sum_{j=1}^{3} z_{ij}\tau_j^2)$, $i = 67...84$ (from (5.8)). Now,

$$E(\delta_i/q_i) = 0, \quad i = 67...84,$$

$$Var(\delta_i/q_i) = (\sum_{j=1}^{3} z_{ij}\tau_j^2)/q_i^2, \quad i = 67...84.$$

Thus, $Var(\delta_i/q_i) < 2.56*10^{-2}$ (Table 5.2) so, with high probability, δ_i/q_i is small compared to 1 and

$$\log(E_i) \approx \log(q_i) + \delta_i/q_i, \quad i = 67...84 \tag{5.13}$$

Hence,

$$\beta_1 E_i + \log(q_i/E_i) \approx \beta_1 q_i + \beta_1 \delta_i - \delta_i/q_i, \quad i = 67...84,$$

and

$$\bar{\beta}_1 \approx \beta_1 + (A' + B' + C' + D')/S[q,q] \tag{5.14}$$

where

$$S[q,q]A' = S[q,q]S[\beta_1\delta - \delta/q,q] - S[\beta_1\delta - \delta/q,q]S[\nu,\nu] -$$

$$S[\beta_1\delta - \delta/q,\nu](2S[q,\nu] + S[\nu,\nu]),$$

$$B' = B,$$

$$C' = C,$$

$$S[q,q]D' = S[q,q]S[\nu/q,\nu] - \beta_1 S[q,q]S[\nu,\nu] -$$

$$(\beta_1 S[q,\nu] + S[\nu/q,q + \nu])(2S[q,\nu] + S[\nu,\nu]).$$

The terms A', B', and C' are linear combinations of $\{\delta_i\}$, $\{\varepsilon_i\}$ and $\{\nu_i\}$ respectively, and D' again contains the remainder of the expression. Now $\{\delta_i\}$ are independent of $\{\varepsilon_i\}$ and $\{\nu_i\}$ so $E(B') = 0$ and $E(C') = 0$ as before. Also, $E(\delta_i) = 0$, $i = 67...84$, so $E(A') = 0$. Thus,

$$E(\hat{\beta}_1) \approx \beta_1 + E(D')/S[q,q]$$

where, using (A5.6), (A5.7), (A5.8) and (A5.9),

$$S[q,q]E(D') = \sum_{j=1}^{3}(\tau_j^2/n_j)\{S[q,q]S[z_j/q - \beta_1 z_j, z_j] -$$

$$2S[q,z_j](S[z_j/q,q] + \beta_1 S[q,z_j]) -$$

$$\sum_{k=1}^{3}(\tau_k^2/n_k)(S[z_j/q,z_j]S[z_k,z_k] + 2S[z_j/q,z_k]S[z_j,z_k])\}.$$

Approximating $\{\mu_j\}$ and $\{\tau_j^2\}$ by their sample estimates $\{\hat{\mu}_j\}$ and $\{\hat{\tau}_j^2\}$ and β_1 by $4.076*10^{-4}$ (its least squares estimate obtained in Chapter 4) gives

$$E(\tilde{\beta}_1 - \beta_1) \approx -7.27*10^{-7}.$$

Since the standard error of $\tilde{\beta}_1$ (estimated ignoring the errors in variables) is $6.32*10^{-6}$, the bias of $\tilde{\beta}_1$ is small.

Approximate expressions for the expectations of $\tilde{\alpha}_1$, $\hat{\alpha}_1$ and $\hat{\beta}_1$ and for the variances of the least squares estimators are at least as complicated as (5.14) and are not derived here.

5.4 APPROXIMATE DISTRIBUTIONS OF THE LEAST SQUARES ESTIMATORS

In this Section, it is assumed that the errors $\{\xi_j\}$ and hence $\{\nu_i\}$ are negligible (ie that $\nu_i = 0$, $i = 67...84$) so that the constants $\{\mu_j\}$ and hence $\{q_i\}$ are known without error. The estimates $\{e_i\}$ are then given by

$$e_i = q_i + \nu_i = q_i, \quad i = 67...84,$$

and are constant. This greatly simplifies the expressions for the least squares estimators - for example, the denominator in (5.9) is constant - so that approximate distributions of $(\tilde{\alpha}_1,\tilde{\beta}_1)'$ and $(\hat{\alpha}_1,\hat{\beta}_1)'$ can be derived.

The errors $\{\nu_i\}$ are negligible if a sufficiently large number of redds are sampled to estimate $\{\mu_j\}$; ie if $\{n_j\}$ are sufficiently large. For Black Brows Beck, the values of $\{n_j\}$ are such that the variance of δ_i is between 3 and 20 times larger then the variance of ν_i, $i = 67...84$ (see Table 5.2). Thus, although the assumption that $\{\nu_i\}$

are negligible is not strictly valid, it is plausible that $\{\delta_i\}$ have a much greater effect on the least squares estimators than $\{\nu_i\}$ and that the distributions derived in this section will be adequate approximations to the true distributions of the least squares estimators. The additional effect of the errors $\{\nu_i\}$ is investigated by some simulations at the end of the Section.

Throughout this Section, the errors $\{\delta_i\}$ are regarded as a realisation of a multivariate normal distribution.

5.4.1 A General Approach

First, consider the linear model

$$y = \eta + \Xi\gamma + \epsilon, \tag{5.15}$$

where y is an n-vector of observations

η is an n-dimensional offset vector

Ξ is an $n \times p$ design matrix

Υ is a p-vector of unknown parameters to be estimated

ϵ is an n-vector of errors distributed $N_n(0, \sigma^2 I_n)$; ϵ is referred to as the observation error.

Suppose that η and Ξ are unobserved, but are approximated by the n-vector a and the $n \times p$ matrix X respectively. Let

$$\eta + \Xi\gamma = a + X\gamma + \zeta. \tag{5.16}$$

The elements of a and X are assumed to be constant (although possibly functions of the (unknown) parameter γ) and the n-vector of "errors" ξ is regarded as a realisation of a multivariate normal distribution with zero mean and variance-covariance matrix Λ. The errors ξ are assumed to be independent of ϵ.

Substituting (5.16) into (5.15) gives

$$y = a + X\gamma + \xi + \epsilon. \tag{5.17}$$

The least squares estimator of γ is given by

$$\hat{\gamma} = (X'X)^{-1}X'(y-a) = \gamma + (X'X)^{-1}X'\epsilon + (X'X)^{-1}X'\xi.$$

The errors ϵ and ξ are independent and normally distributed with zero means. Also, X is constant. Hence, $\hat{\gamma}$ is normally distributed with

$$E(\hat{\gamma}) = \gamma,$$

$$\mathrm{Var}(\hat{\gamma}) = (X'X)^{-1}\sigma^2 + (X'X)^{-1}X'\Lambda X(X'X)^{-1}. \tag{5.18}$$

Thus the least squares estimator is unbiased. Also, the variance of $\hat{\gamma}$ can be partitioned into the variance due to the errors in the observation y, namely $(X'X)^{-1}\sigma^2$, and the variance due to the errors-in-variables, namely $(X'X)^{-1}X'\Lambda X(X'X)^{-1}$.

5.4.2 Multiplicative Error Model

Since $e_i = q_i$, $i = 67...84$, (5.4) and (5.13) give

$$\log(e_i + \delta_i) \approx \log(e_i) + \delta_i/e_i, \quad i = 67...84.$$

Hence, the multiplicative error Ricker model (5.7b) can be approximated by

$$\log(R_{1i}) = \log(e_i) + \alpha_1 - \beta_1 e_i + \delta_i(\frac{1}{e_i} - \beta_1) + \epsilon_i, \quad i = 67...84.$$

This is of the same form as (5.17) with

$$y_i = \log(R_{1i}), \quad i = 67...84,$$

$$\alpha_i = \log(e_i), \quad i = 67...84$$

$$X = \begin{pmatrix} 1,...,1 \\ -e_{67},...,-e_{84} \end{pmatrix}' \tag{5.19}$$

$$\gamma = (\alpha_1, \beta_1)',$$

$$\xi_i = \delta_i(\frac{1}{e_i} - \beta_1), \quad i = 67...84,$$

$$\Lambda = \text{diag}(\tau_{67}^2(\frac{1}{e_{67}} - \beta_1)^2,..., \tau_{84}^2(\frac{1}{e_{84}} - \beta_1)^2). \tag{5.20}$$

Also, the matrix X is constant, since $\{e_i\}$ are constant. Thus, the results of Section 5.4.1 can be applied. The least squares estimator $(\tilde{\alpha}_1, \tilde{\beta}_1)'$ is approximately normally distributed, unbiased, with a variance given by (5.18), (5.19) and (5.20).

It is interesting to note from the form of (5.20) that errors in both low and high egg values have a much greater effect on the variance of the least squares estimator than errors in medium egg values (ie values close to $1/\beta_1$).

The effect of the errors $\{\delta_i\}$ is seen by considering the case $\alpha_1 = -0.7291$, $\beta_1 = 4.076*10^{-4}$, $\sigma^2 = 4.085*10^{-3}$ (the least squares estimates obtained in Chapter 4). Two 95% contours are shown in Figure 5.1. The inner contour defines a region within which the least squares estimator will lie with probability 0.95 if there are no errors-in-variables (ie if $\Lambda = 0$). The outer contour defines the corresponding region when Λ is given by (5.20). The errors $\{\delta_i\}$ considerably enlarge the 95% contour.

The distribution of $(\tilde{\alpha}_1,\tilde{\beta}_1)'$ is also investigated by a simulation. One thousand sets of recruits are simulated using (5.7b) with the parameter values above, by generating 1,000 sets of random errors $\{\varepsilon_i\}$ and $\{\delta_i\}$ (but taking $\nu_i = 0$, $i = 67...84$). The least squares estimates obtained from the simulated data sets are shown in Figure 5.1 and indicate that the effect of the errors $\{\delta_i\}$ is well approximated by the distribution above since 45 of the 1000 points (ie approximately 5%) lie outside the outer 95% contour.

A similar simulation is carried out in which all three sets of errors $\{\xi_j\}$ (and hence $\{\nu_i\}$), $\{\varepsilon_i\}$ and $\{\delta_i\}$ are generated. The corresponding least squares estimates are shown in Figure 5.2 with the outer 95% contour from Figure 5.1. The errors $\{\nu_i\}$ increase the variance of the least squares estimator even more. However, the distribution of $(\tilde{\alpha}_1,\tilde{\beta}_1)'$ obtained above is still a considerable improvement on the distribution used if the errors-in-variables are ignored.

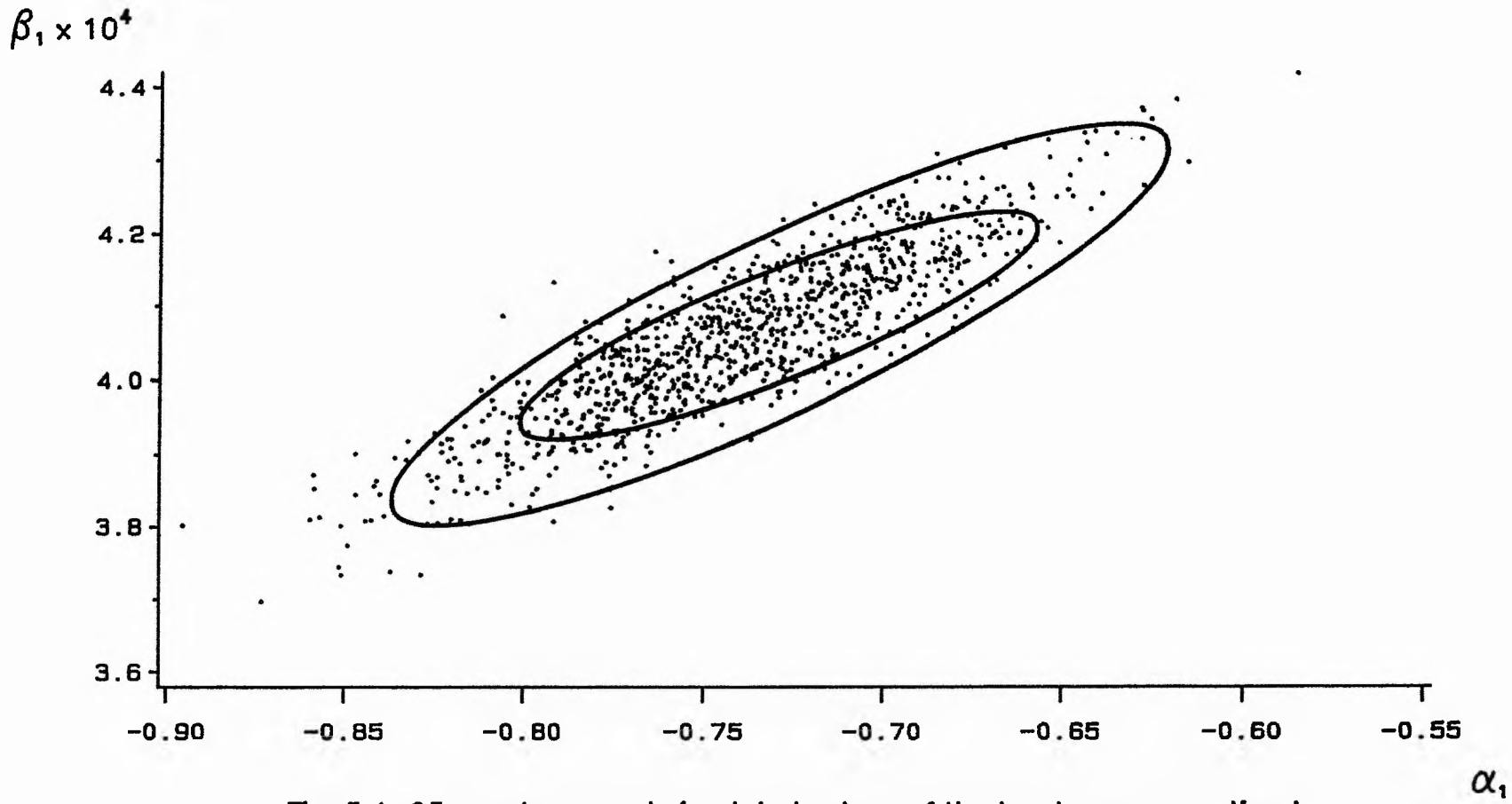


Fig. 5.1 95% contours and simulated values of the least squares estimator of the multiplicative error Ricker model; the inner contour ignores the errors–in–variables; the outer contour and the simulated values incorporate the errors $|\delta|$ (but ignore the errors $|\nu|$).

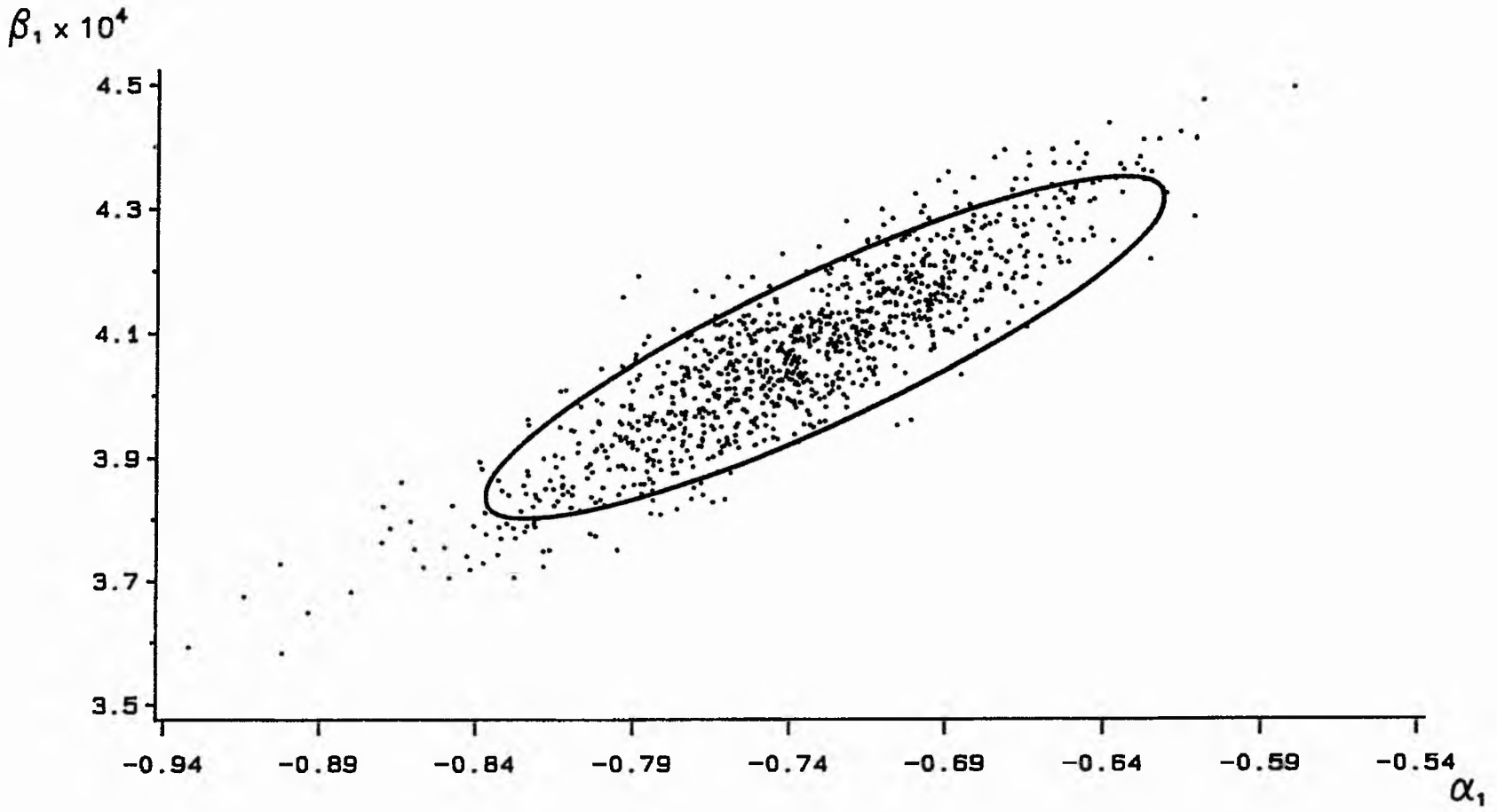


Fig. 5.2 95% contour and simulated values of the least squares estimator of the multiplicative error Ricker model; the contour incorporates the errors $\{\delta\}$ but ignores the errors $\{\nu\}$; the simulated values incorporate both $\{\delta\}$ and $\{\nu\}$.

5.4.3 Additive Error Model

Let α_{10} and β_{10} be the true values of α_1 and β_1. The additive error Ricker model (5.7a) can be written

$$\begin{aligned} R_{1i} &= (e_i + \delta_i)\exp(\alpha_{10} - \beta_{10}(e_i + \delta_i)) + \epsilon_i \\ &= e_i\exp(\alpha_{10} - \beta_{10} e_i)(1 + \delta_i/e_i)\exp(-\beta_{10}\delta_i) + \epsilon_i \\ &= h_i(\alpha_{10},\beta_{10})(1 + \delta_i/e_i)\exp(-\beta_{10}\delta_i) + \epsilon_i, \quad i = 67...84, \end{aligned} \tag{5.21}$$

where

$$h_i(\alpha,\beta) = e_i\exp(\alpha - \beta e_i), \quad i = 67...84.$$

Now, δ_i/e_i is small compared to 1 with high probability, $i = 67...84$ (Section 5.3). Also $\beta_{10}\delta_i$ is small compared to 1 with high probability since $E(\beta_{10}\delta_i) = 0$ and $Var(\beta_{10}\delta_i) < 0.026$, $i = 67...84$ (Table 5.2), (where β_{10} is approximated by its least squaresestimate obtained in Chapter 4). Thus, (5.21) can be approximated by

$$R_{1i} \cong h_i(\alpha_{10},\beta_{10}) + \delta_i h_i(\alpha_{10},\beta_{10})\left(\frac{1}{e_i} - \beta_{10}\right) + \epsilon_i, \quad i = 67...84. \tag{5.22}$$

For values of α_{10}, β_{10} and σ^2 such as those obtained in Chapter 4, the nonlinearity of the Ricker model is small in a large region of interest around (α_{10},β_{10}) (Chapter 4). Thus, (5.22) can be well approximated by a linearisation about the least squares estimator $(\hat{\alpha}_1,\hat{\beta}_1)$

$$R_{1i} \approx h_i(\hat{\alpha}_1,\beta_1) + (\alpha_{10} - \hat{\alpha}_1) \left.\frac{dh_i}{d\alpha_1}\right|_{(\hat{\alpha}_1,\beta_1)} + (\beta_{10} - \beta_1) \left.\frac{dh_i}{d\beta_1}\right|_{(\hat{\alpha}_1,\beta_1)}$$

$$+ \delta_i h_i(\alpha_{10},\beta_{10}) \left(\frac{1}{e_i} - \beta_{10}\right) + \epsilon_i$$

$$\approx h_i(\hat{\alpha}_1,\beta_1) + (\alpha_{10} - \hat{\alpha}_1) \left.\frac{dh_i}{d\alpha_1}\right|_{(\alpha_{10},\beta_{10})} + (\beta_{10} - \beta_1) \left.\frac{dh_i}{d\beta_1}\right|_{(\alpha_{10},\beta_{10})}$$

$$+ \delta_i h_i(\alpha_{10},\beta_{10}) \left(\frac{1}{e_i} - \beta_{10}\right) + \epsilon_i, \quad i = 67...84.$$

This equation is of the same form as (5.17) with

$$X = \begin{pmatrix} \left.\frac{dh_{67}}{d\alpha_1}\right|_{(\alpha_{10},\beta_{10})} & \cdots & \left.\frac{dh_{84}}{d\alpha_1}\right|_{(\alpha_{10},\beta_{10})} \\ \left.\frac{dh_{67}}{d\beta_1}\right|_{(\alpha_{10},\beta_{10})} & \cdots & \left.\frac{dh_{84}}{d\beta_1}\right|_{(\alpha_{10},\beta_{10})} \end{pmatrix}' \tag{5.23}$$

$$\Lambda = \mathrm{diag}(\tau_{67}^2 h_{67}^2(\alpha_{10},\beta_{10})(\frac{1}{e_{67}} - \beta_{10})^2, \ldots, \tau_{84}^2 h_{84}^2(\alpha_{10},\beta_{10})(\frac{1}{e_{84}} - \beta_{10})^2). \tag{5.24}$$

Again, X is constant. Thus the least squares estimator $(\hat{\alpha}_1,\hat{\beta}_1)'$ is approximately normally distributed, unbiased, with a variance given by (5.18), (5.23) and (5.24). The 95% contours and the results of the simulations for the additive error model with the true parameter values $\alpha_{10} = -0.7392$, $\beta_{10} = 4.065*10^{-4}$, $\sigma^2 = 347.2$ are shown in Figures 5.3 and 5.4. Again, both errors $\{\delta_i\}$ and $\{\nu_i\}$ considerably increase the variance of the least squares estimator. Also, the effect of the errors $\{\delta_i\}$ is well approximated by the distribution derived above.

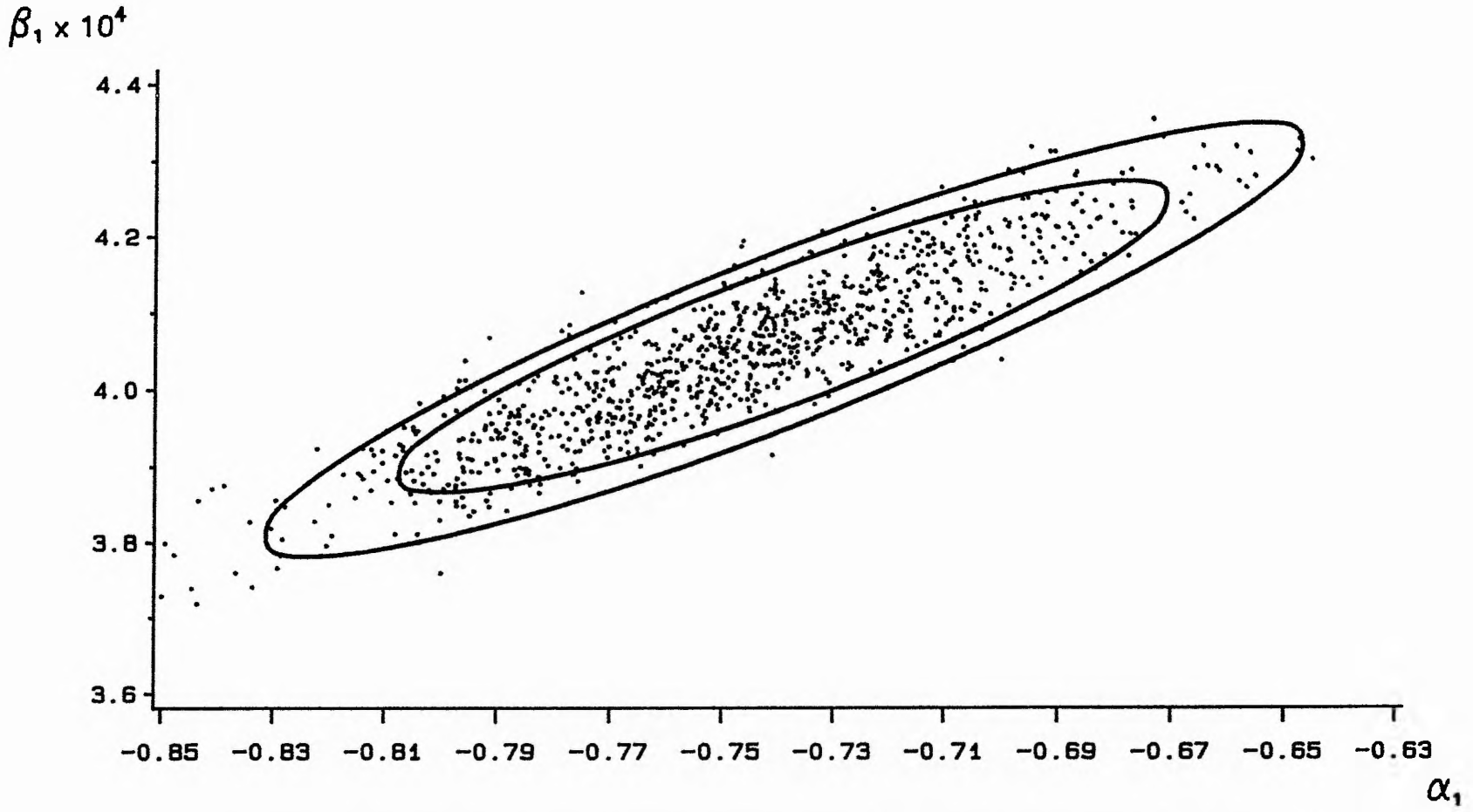


Fig. 5.3 95% contours and simulated values of the least squares estimator of the additive error Ricker model; the inner contour ignores the errors–in–variables; the outer contour and the simulated values incorporate the errors $\{\delta\}$ (but ignore the errors $\{\nu\}$).

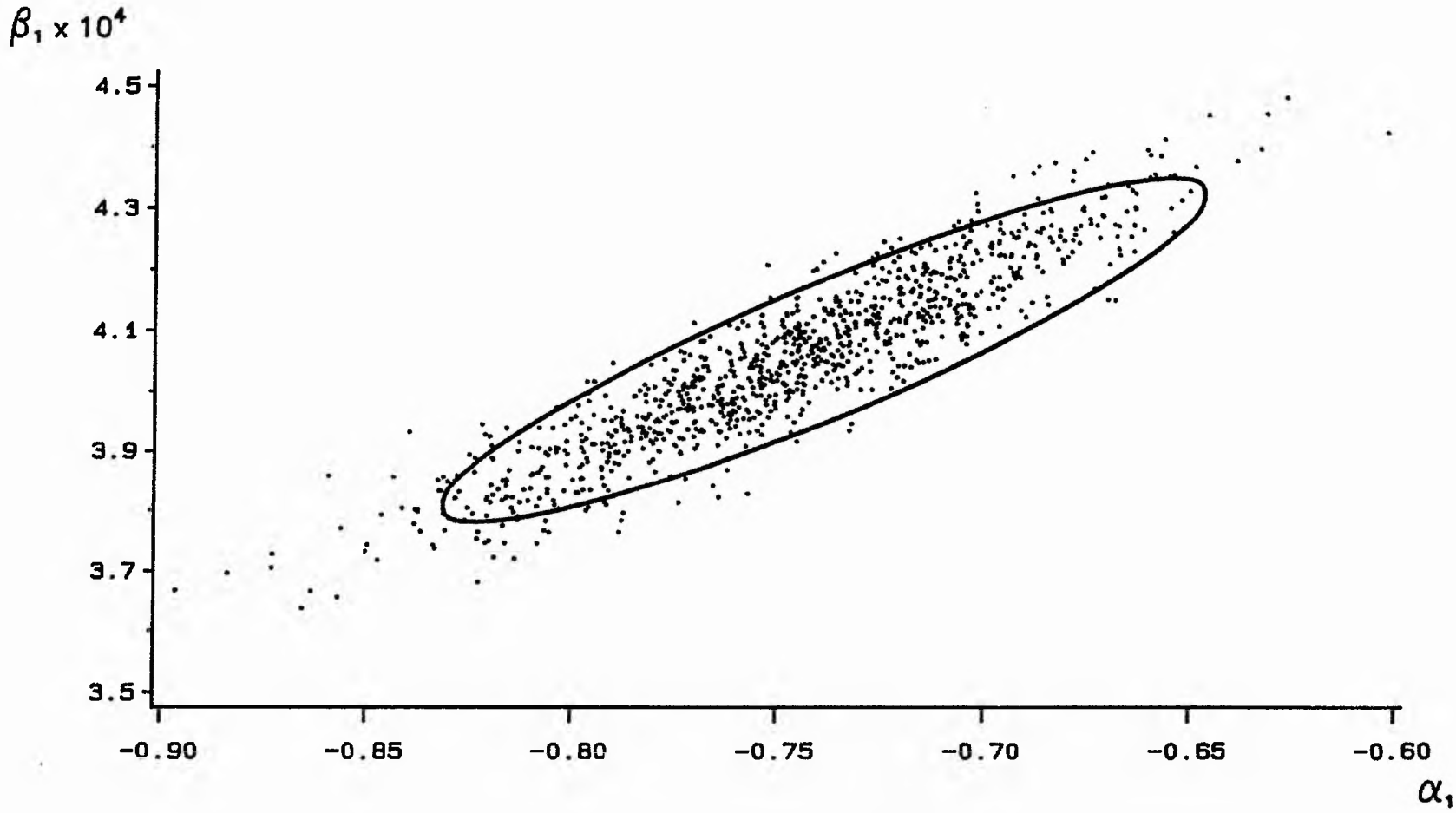


Fig. 5.4 95% contour and simulated values of the least squares estimator of the additive error Ricker model; the contour incorporates the errors $\{\delta\}$ but ignores the errors $\{\nu\}$; the simulated values incorporate both $\{\delta\}$ and $\{\nu\}$.

5.5 DISCUSSION

The errors-in-variables pose many problems of inference. If σ^2 is known, the results of Section 5.4 can be used to account for the effect of the errors $\{\delta_i\}$. For example, improved approximate 95% confidence regions for $(\alpha_1,\beta_1)'$ are obtained by using the approximate distributions of $(\tilde{\alpha}_1,\tilde{\beta}_1)'$ or $(\hat{\alpha}_1,\hat{\beta}_1)'$ derived above. However, the effect of the errors $\{\nu_i\}$ is still not taken into account.

Inference is much more complicated if σ^2 is unknown, as for the Black Brows Beck data. For example, consider the estimation of σ^2. The usual estimator s^2, based on the residual sum of squares, can greatly overestimate σ^2 if the errors-in-variables are "large". This is because s^2 incorporates the variation due to the errors-in-variables as well as the variation due to the observation errors. For example, in the case of the linear model (5.15),

$$s^2 = \frac{1}{(n-p)}(y-\hat{y})'(y-\hat{y})$$

$$= \frac{1}{(n-p)}(\epsilon+\zeta)'(I-X(X'X)^{-1}X')(\epsilon+\zeta),$$

so that

$$E(s^2) = \sigma^2 + \frac{1}{(n-p)}E(\zeta'(I-X(X'X)^{-1}X')\zeta)$$

$$> \sigma^2$$

since $I - X(X'X)^{-1}X'$ is positive definite. Assuming Λ is known and diagonal, an unbiased estimator of σ^2 (in the linear model (5.15)) is

$$s_1^2 = s^2 - \frac{1}{(n-p)} E(\zeta'(I - X(X'X)^{-1}X')\zeta)$$

$$= s^2 - \frac{1}{(n-p)} \sum_{i=1}^{n} \Lambda_{ii}(I - X(X'X)^{-1}X')_{ii}$$

however, this is also unsatisfactory, since it can give negative estimates of σ^2. Thus, an alternative approach is required if satisfactory estimates of σ^2 are to be obtained.

A detailed study of the problems of inference caused by the errors-in-variables is beyond the scope of this Thesis. Instead, the following approach is adopted. It is assumed that the errors-in-variables models (5.7) can be adequately approximated by the models

$$R_{1i} = e_i \exp(\alpha_1 - \beta_1 e_i) + \epsilon_i^+, \quad i = 67...84, \tag{5.25a}$$

$$\log(R_{1i}) = \log(e_i) + \alpha_1 - \beta_1 e_i + \epsilon_i^+, \quad i = 67...84, \tag{5.25b}$$

where $\{\epsilon_i^+\}$ are NID$(0,(\sigma^2)^+)$. The errors $\{\epsilon_i^+\}$ incorporate both the observation errors and the errors-in-variables. Standard regression theory is then used to obtain parameter estimates, confidence regions etc. Since an estimate of $(\sigma^2)^+$ based on the residual sum of squares includes variation due to the errors-in-variables, it is assumed that these confidence regions are adequate approximations to exact confidence regions etc. The estimates and confidence regions obtained in this way are identical to those obtained if the errors-in-variables problem is not considered at all. However, there is an important difference in interpretation. If the errors-in-variables are ignored, the Ricker models (5.25) are interpreted as stock-recruitment models relating R_1 to the numbers of eggs. If the

errors-in-variables are considered, models (5.25) are stock-recruitment models relating R_1 to the numbers of redds in each of the three spawning periods, since (5.25) can be written as

$$R_{1i} = (\sum_{j=1}^{3} z_{ij}\hat{\mu}_j)\exp(\alpha_1 - \beta_1 \sum_{j=1}^{3} z_{ij}\hat{\mu}_j) + \epsilon_i^*, \quad i = 67...84,$$

$$\log(R_{1i}) = \log(\sum_{j1}^{3} z_{ij}\hat{\mu}_j) + \alpha_1 - \beta_1 \sum_{j=1}^{3} z_{ij}\hat{\mu}_j + \epsilon_i^*, \quad i = 67...84.$$

The errors-in-variables problem is not considered further. In subsequent Chapters, although the errors-in-variables are implicitly acknowledged to exist, they are effectively ignored.

APPENDIX 5

Let ν be the vector of errors

$$\nu = (\nu_{67}, \ldots, \nu_{84})',$$

where $\nu_i = \sum_{j=1}^{3} z_{ij}\xi_j$, $i = 67...84$, and where $\{\xi_j\}$ are independent with $\xi_j \sim N(0, \tau_j^2/n_j)$, $j = 1...3$. Let

$$z_j = (z_{67j}, \ldots, z_{84j})', \qquad j = 1...3.$$

and let ψ, θ and ϕ be any 18-dimensional vectors that are independent of ν. For notational convenience, denote the 18-dimensional vector with ith element $(\psi_i\theta_i)$, $i = 67...84$, by $\psi * \theta$.

i)
$$S[\nu, \nu] = \sum_{i=67}^{84} (\nu_i - \bar{\nu})(\nu_i - \bar{\nu})$$

$$= \sum_{i=67}^{84} \{\sum_{j=1}^{3} \xi_j (z_{ij} - \bar{z}_j)\}\{\sum_{k=1}^{3} \xi_k (z_{ik} - \bar{z}_k)\}$$

$$= \sum_{j=1}^{3}\sum_{k=1}^{3} \xi_j\xi_k S[z_j, z_k]. \qquad (A5.1)$$

Similarly,

$$S[q, \nu] = \sum_{j=1}^{3} \xi_j S[q, z_j] \tag{A5.2}$$

$$S[\psi^* \nu, \nu] = \sum_{j=1}^{3} \sum_{k=1}^{3} \xi_j \xi_k S[\psi^* z_j, z_k]. \tag{A5.3}$$

ii) From (A5.2) and because ξ_j, $j = 1...3$, are independent with $\xi_j \sim N(0, \tau_j^2/n_j)$

$$E(S[q, \nu]) = 0 \tag{A5.4}$$

$$Var(S[q, \nu]) = \sum_{j=1}^{3} (\tau_j^2/n_j)(S[q, z_j])^2. \tag{A5.5}$$

iii)
$$E(S[\psi^* \nu, \nu]) = E(\sum_{j=1}^{3} \sum_{k=1}^{3} \xi_j \xi_k S[\psi^* z_j, z_k])$$

$$= \sum_{j=1}^{3} (E(\xi_j^2)) S[\psi^* z_j, z_j]$$

$$= \sum_{j=1}^{3} (\tau_j^2/n_j) S[\psi^* z_j, z_j]. \tag{A5.6}$$

iv) $E(S[\psi^* \nu, \theta] S[\nu, \phi])$

$$= E(\sum_{j=1}^{3} \sum_{k=1}^{3} \xi_j \xi_k S[\psi^* z_j, \theta] S[z_k, \phi])$$

$$= \sum_{j=1}^{3} (\tau_j^2/n_j) S[\psi^* z_j, \theta] S[z_j, \phi]. \tag{A5.7}$$

v) $E(S[\psi^* v, v]S[\theta^* v, \phi])$

$$= E\Big(\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\xi_j\xi_k\xi_l S[\psi^* z_j, z_k]S[\theta^* z_l, \phi]\Big) = 0. \qquad \text{(A5.8)}$$

vi) $E(S[\psi^* v, v]S[v, v])$

$$= E\Big(\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\sum_{m=1}^{3}\xi_j\xi_k\xi_l\xi_m \; S[\psi^* z_j, z_k]S[z_l, z_m]\Big)$$

$$= \sum_{j=1}^{3} 3(\tau_j^2/n_j)^2 \; S[\psi^* z_j, z_j] \; S[z_j, z_j] +$$

$$\sum_{j=1}^{3}\sum_{k=1}^{3}{}_{k\neq j}(\tau_j^2/n_j)(\tau_k^2/n_k)S[\psi^* z_j, z_j]S[z_k, z_k] +$$

$$2\sum_{j=1}^{3}\sum_{k=1}^{3}{}_{k\neq j}(\tau_j^2/n_j)(\tau_k^2/n_k)S[\psi^* z_j, z_k]S[z_j, z_k]$$

$$= \sum_{j=1}^{3}\sum_{k=1}^{3}(\tau_j^2/n_j)(\tau_k^2/n_k)(S[\psi^* z_j, z_j]S[z_k, z_k] +$$

$$2S[\psi^* z_j, z_k]S[z_j, z_k]). \qquad \text{(A5.9)}$$

CHAPTER 6 DROUGHT AND SURVIVAL

6.1 INTRODUCTION

Summer drought is the only environmental factor found to have an important effect on the survival of trout in Black Brows Beck (Elliott, 1985a). High mortality rates during droughts are mainly due to the low water level of the beck, which greatly restricts the trout's habitat. For example, in 1976, 1983 and 1984, large parts of the beck dried up completely (Elliott, 1985a). Other possible causes of mortality are a reduction in the food supply, a reduction in the water velocity of the beck and an increase in the water temperature.

The life stages R_2, R_3 and R_4 can all be affected by periods of summer drought (where R_2, R_3 and R_4 are the number of 0+ parr at time t_2 (the August/September electro-fishing), the number of 1+ parr at time t_3 (the May/June electro-fishing) and the number of 1+ parr at time t_4 (the August/September electro-fishing) respectively (Chapter 1)). All three stages can be affected by droughts in the first summer of the life cycle (ie between t_1 and t_2). In addition, R_4 can be affected by droughts in the second summer of the life cycle (ie between t_3 and t_4). Simple stock-recruitment models, such as the Ricker model (2.6a), can not adequately describe the relationships between R_2, R_3, R_4 and E, the number of eggs, because they fail to account for the effect of summer drought on survival.

This Chapter investigates the relationship between drought and survival in more detail. A measure of drought based on the water level of Esthwaite Water, a lake close to Black Brows Beck, is developed in Section 6.2. This measure is then used

in the development of a model that gives a good description of the relationship between R_2 and E, even in years affected by drought. The model is extended to describe the relationships between R_3 and E and between R_4 and E in Sections 6.3 and 6.4 respectively. Finally, the relationship between E^*, the egg production of a year class, and E is investigated in Section 6.5.

6.2 THE RELATIONSHIP BETWEEN R_2 AND E

6.2.1 Drought Years

When the additive error Ricker model

$$R_{2i} = E_i \exp(\alpha_2 - \beta_2 E_i) + \epsilon_i, \qquad i = 67 \ldots 84 \tag{6.1}$$

where $\{\epsilon_i\}$ are NID$(0, \sigma^2)$, is fitted to the R_2 data, two large negative residuals are identified (Fig. 6.1). These correspond to the years 1983 and 1984, in which there were periods of summer drought between t_1 and t_2. If the 1983 and 1984 data points are omitted and the Ricker model is fitted again, three new large negative residuals appear; these correspond to the years 1968, 1969, 1976, in which there were also summer droughts (but not as severe droughts as those of 1983 and 1984).

If the 1968, 1969, 1976, 1983 and 1984 data points are omitted, the additive error Ricker model is an excellent description of the relationship between R_2 and E (Fig. 6.2). An examination of the residuals gives no evidence that the choice of error structure is inappropriate. Parameter estimates are given in Table 6.1. Thus, it is assumed that drought had no effect on survival in any of the remaining years and

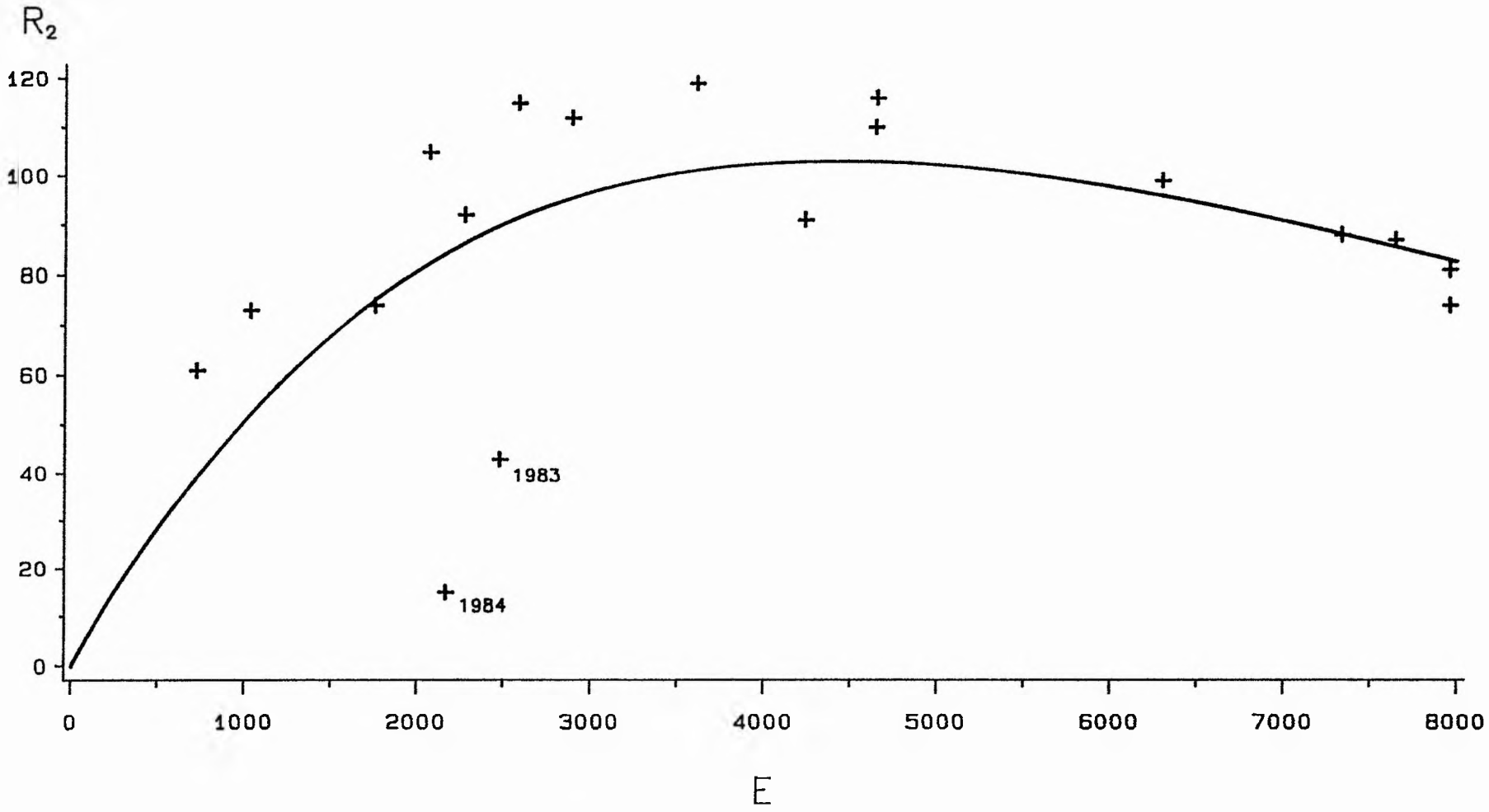


Fig. 6.1 A Ricker curve fitted to the R_2 data showing the large negative residuals in the drought years 1983 and 1984.

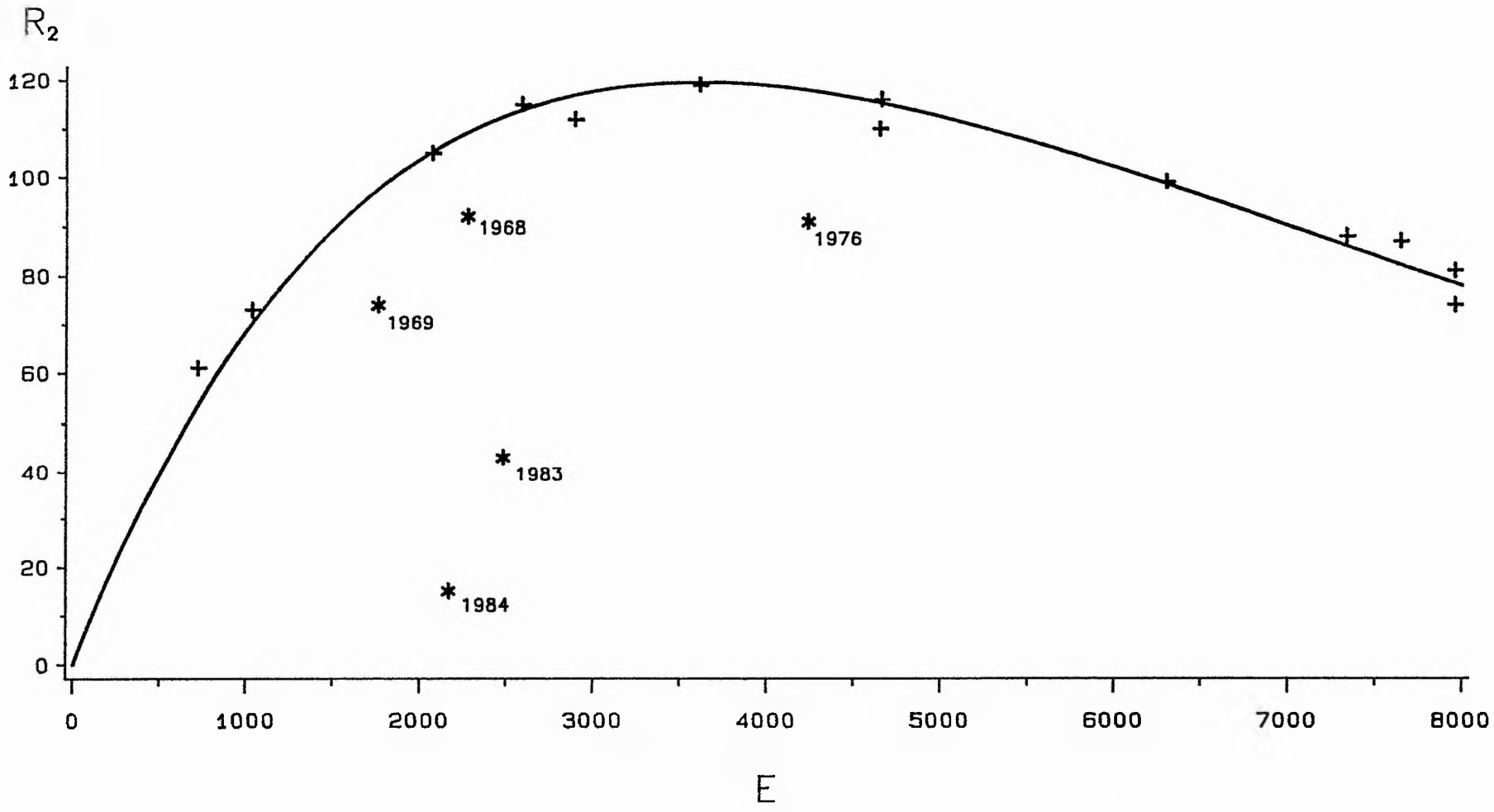


Fig. 6.2 A Ricker curve fitted to the R_2 data excluding the drought years. The drought years (*) are shown for comparison.

TABLE 6.1

Parameter estimates for model (6.1) (drought years omitted)

Parameter	Estimate	Lin approx s.e.	Correlations	
α_2	- 2.400	0.026	1.000	
β_2	$2.790*10^{-4}$	$0.053*10^{-4}$	0.901	1.000

rss = 172.35 on 11 d.f.
s^2 = 15.67

that in the absence of drought, the relationship between R_2 and E is well described by (6.1).

The years 1968, 1969, 1976, 1983 and 1984 are collectively referred to as the "drought years".

6.2.2 Water Level and Rainfall Data

There are no detailed long term records of the water levels, water velocities etc of Black Brows Beck. Therefore, the "drought effect" is investigated by using water level and rainfall figures from two sites close to Black Brows Beck. Two data sets are used:

i) the daily water level of Esthwaite Water (1967-84),

ii) the daily rainfall at Ambleside (1967-84).

Since the water level is the main factor affecting survival during drought, the water level data are used first to identify patterns of water levels common to the drought years and hence to construct a measure of drought. Rainfall figures are less useful because the water level, water velocity etc of the beck on a given day depend in some unknown way on the rainfall (and temperature, humidity etc) on the preceding days. However, rainfall data are all that is available in many situations, so the relationship between the Ambleside rainfall and survival is investigated in Section 6.2.6.

6.2.3 Critical Period for Survival

The mean water level of Esthwaite changed dramatically due to drainage work twice between 1967 and 1984; namely during the winters of 1973-74 and 1980-81. Thus,

the water level data are studied by partitioning the years 1967-84 into the three periods, 1967-73, 1974-80 and 1981-84. The water levels between t_1 and t_2 for these periods are shown in Figure 6.3, with the water levels in the drought years plotted in a continuous line. Two main features emerge. First, there are long periods of low water level in almost every year. Secondly, the water level in the drought years is always low in the period immediately before t_2. Thus, the timing of a period of low water level is an important factor in determining whether the period of low water level has a marked effect on survival.

It is useful to consider the water levels in the drought year 1976 and the non-drought year 1978 in more detail (Fig. 6.4). The water level in 1978 is lower than the water level in 1976 throughout almost all the period 1 June - 13 August. Thus, low water levels in a particular calendar period (eg during June and July) do not necessarily reduce the survival rate between t_1 and t_2.

However, an important difference between 1976 and 1978 is the sampling date t_2; namely, 13 August in 1976 and 30 August in 1978. Each year, t_2 is approximately 125 days after the emergence of the alevins (Elliott, 1984b). Thus, regardless of the year being considered, the trout at t_2 are all at approximately the same stage of the life cycle. Thus, it is sensible to consider water levels relative to the time t_2. In particular, the water level is low in the 30 days before t_2 in the drought year 1976, but high in the 30 days before t_2 in the non-drought year 1978. Plotting the water level against time relative to t_2 shows that this pattern occurs throughout the period 1967-84; in the drought years, the water level is always low in the 30 days before t_2; in the non-drought years, there is always at least one period of high water level in the 30 days before t_2 (Fig. 6.5). Thus, the survival rate is reduced if the

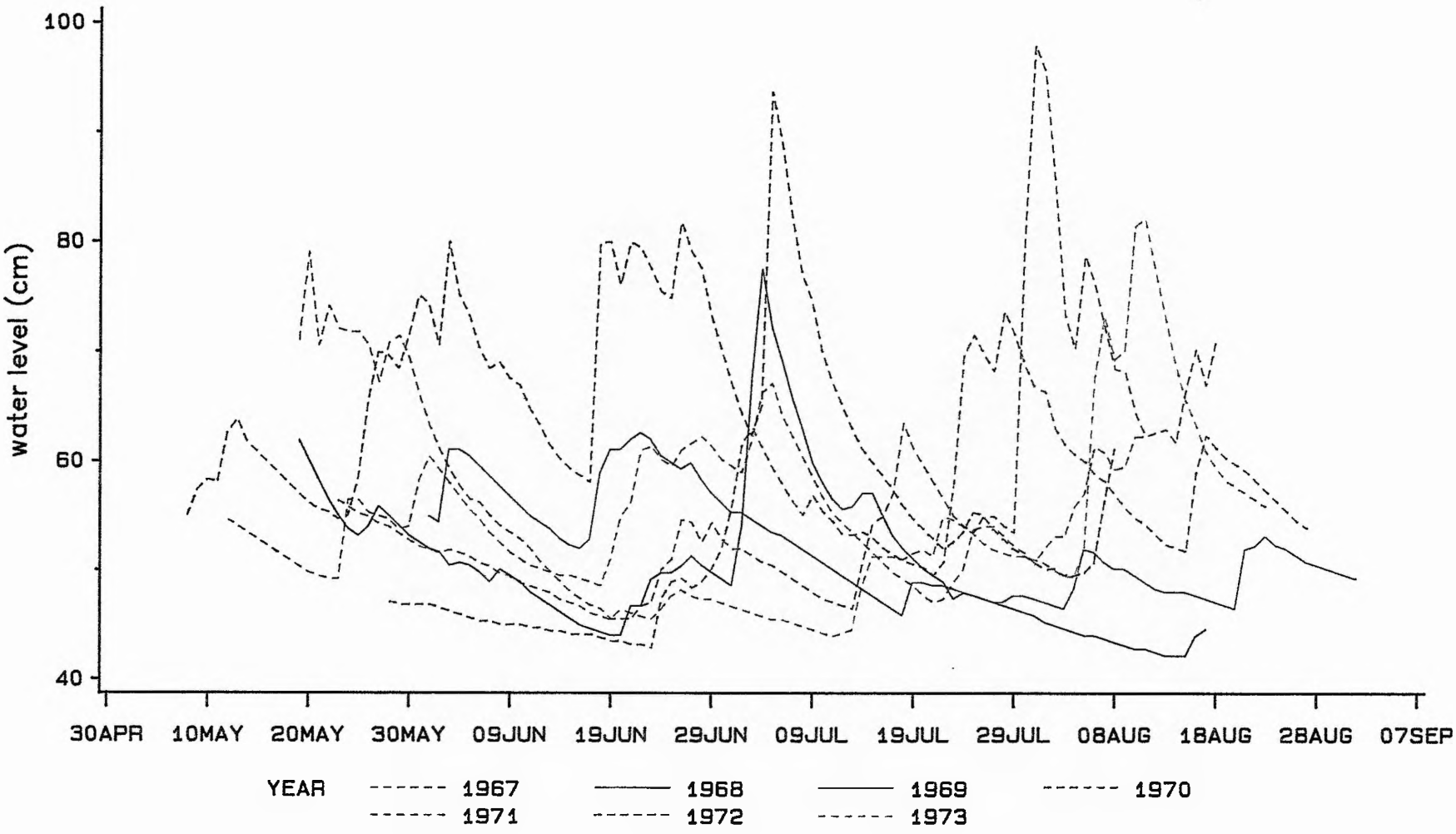


Fig. 6.3a Water level between t_1 and t_2 1967 – 1973.

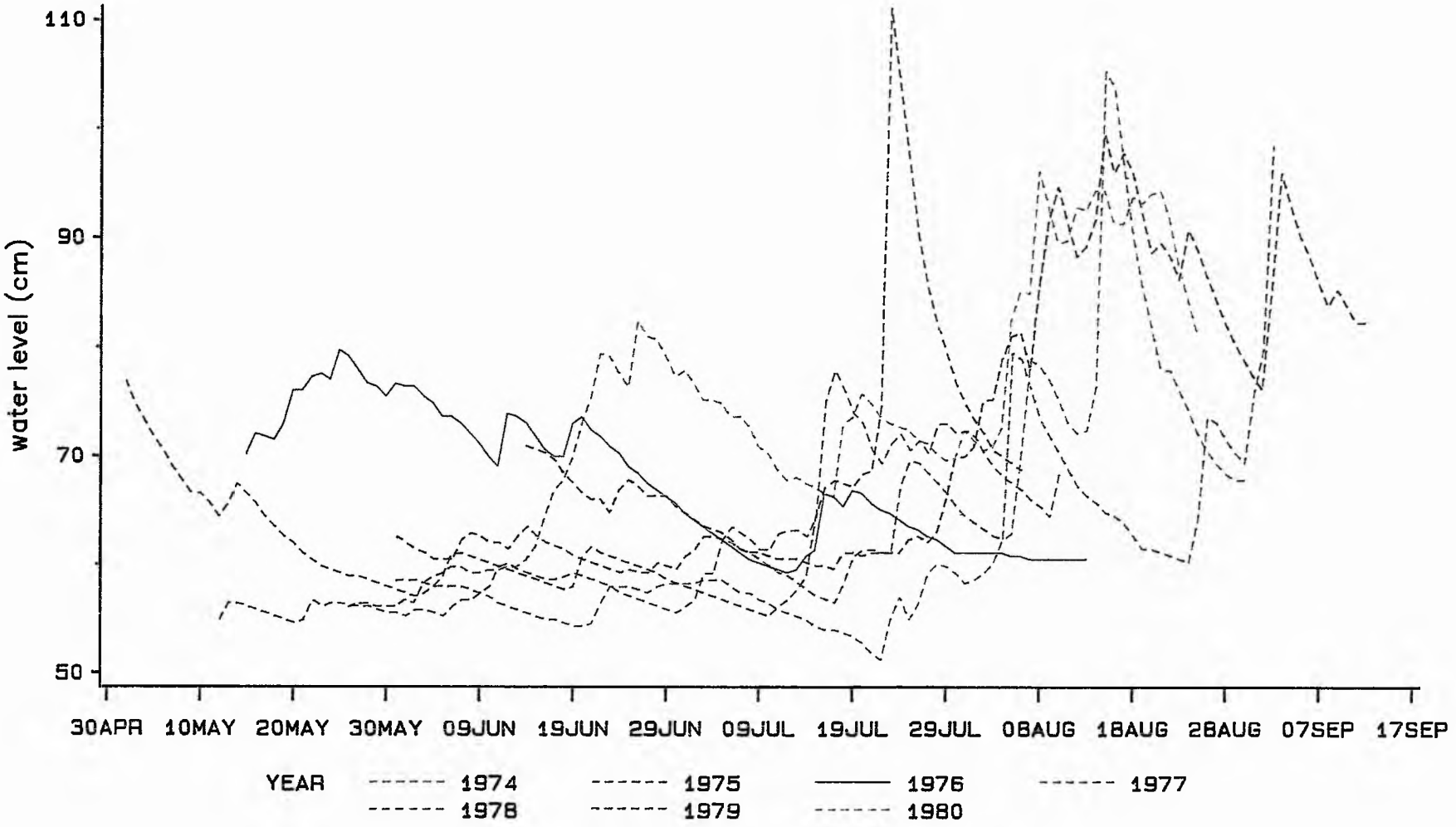


Fig. 6.3b Water level between t_1 and t_2 1974 – 1980.

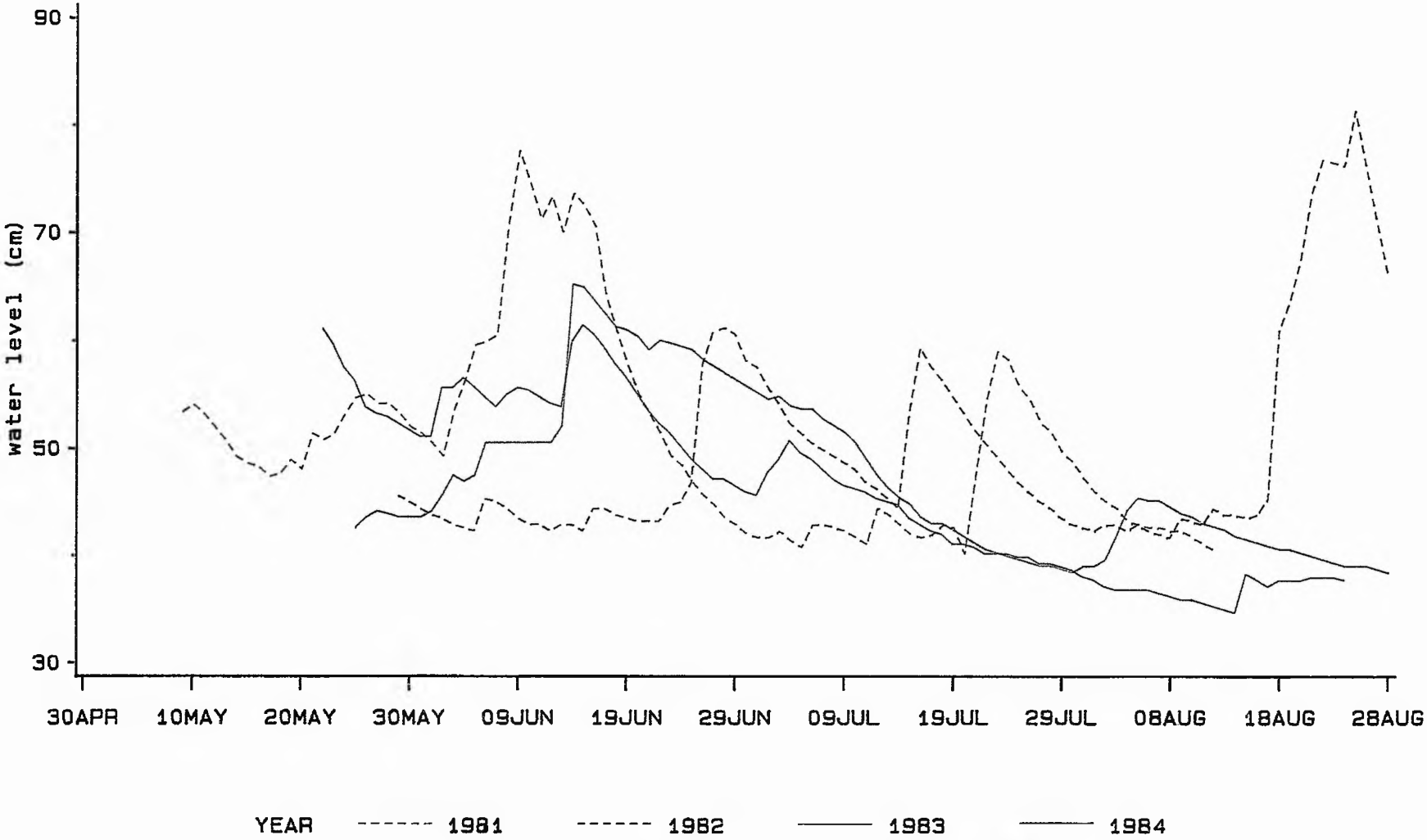


Fig. 6.3c Water level between t_1 and t_2 1981 - 1984.

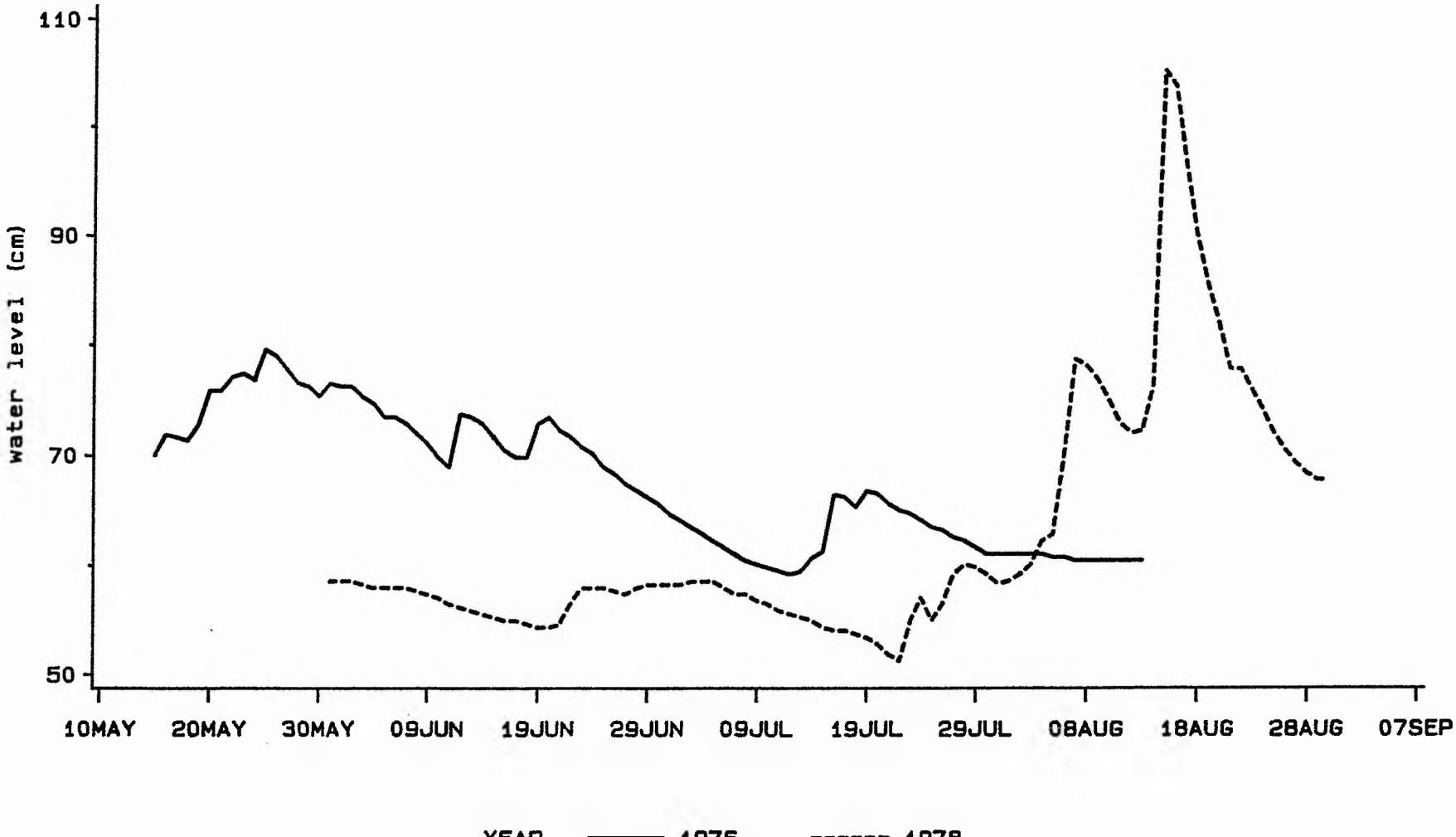


Fig. 6.4 Water level between t_1 and t_2 1976 and 1978.

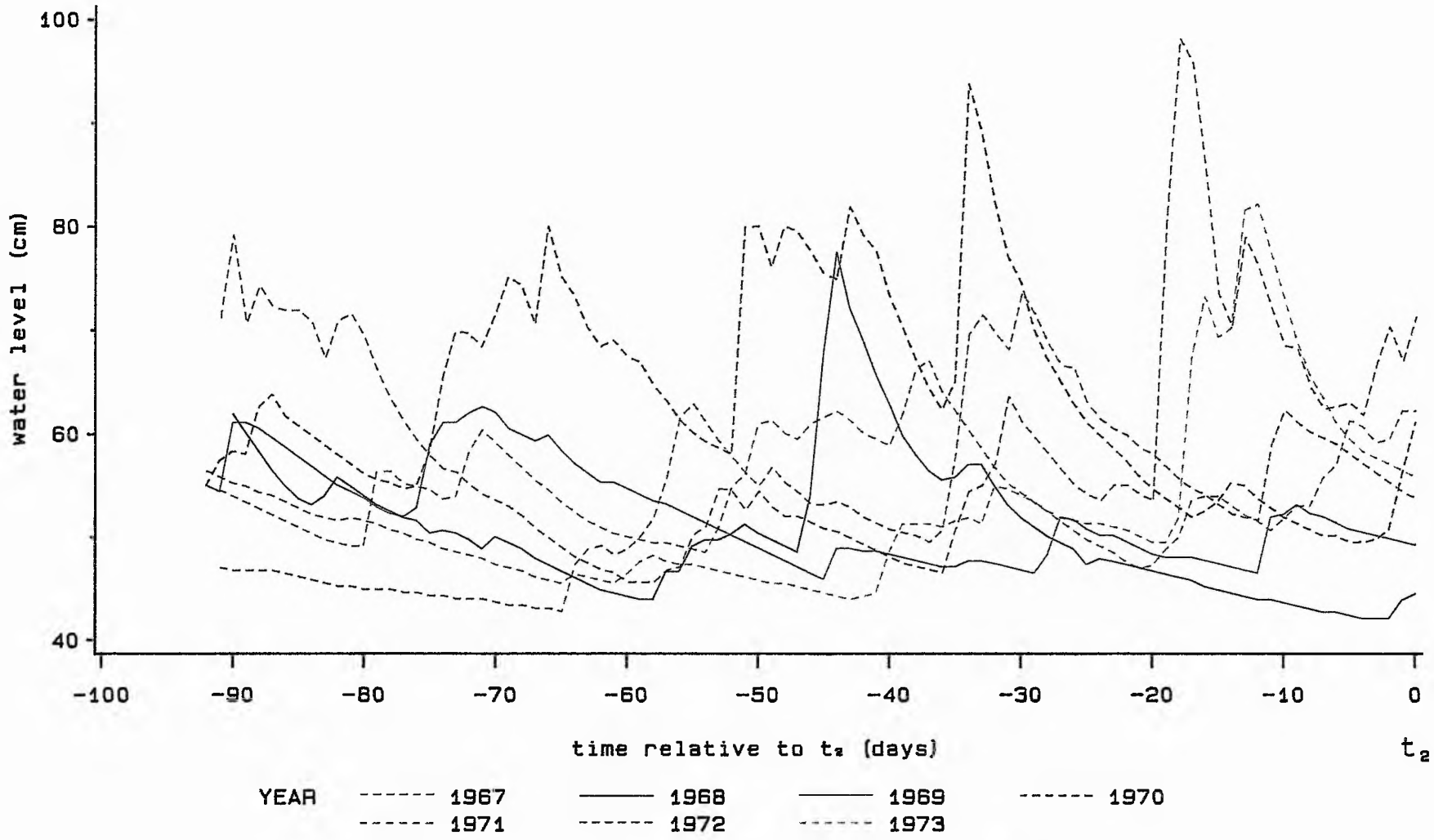


Fig. 6.5a Water level vs. time relative to t_2 1967 - 1973.

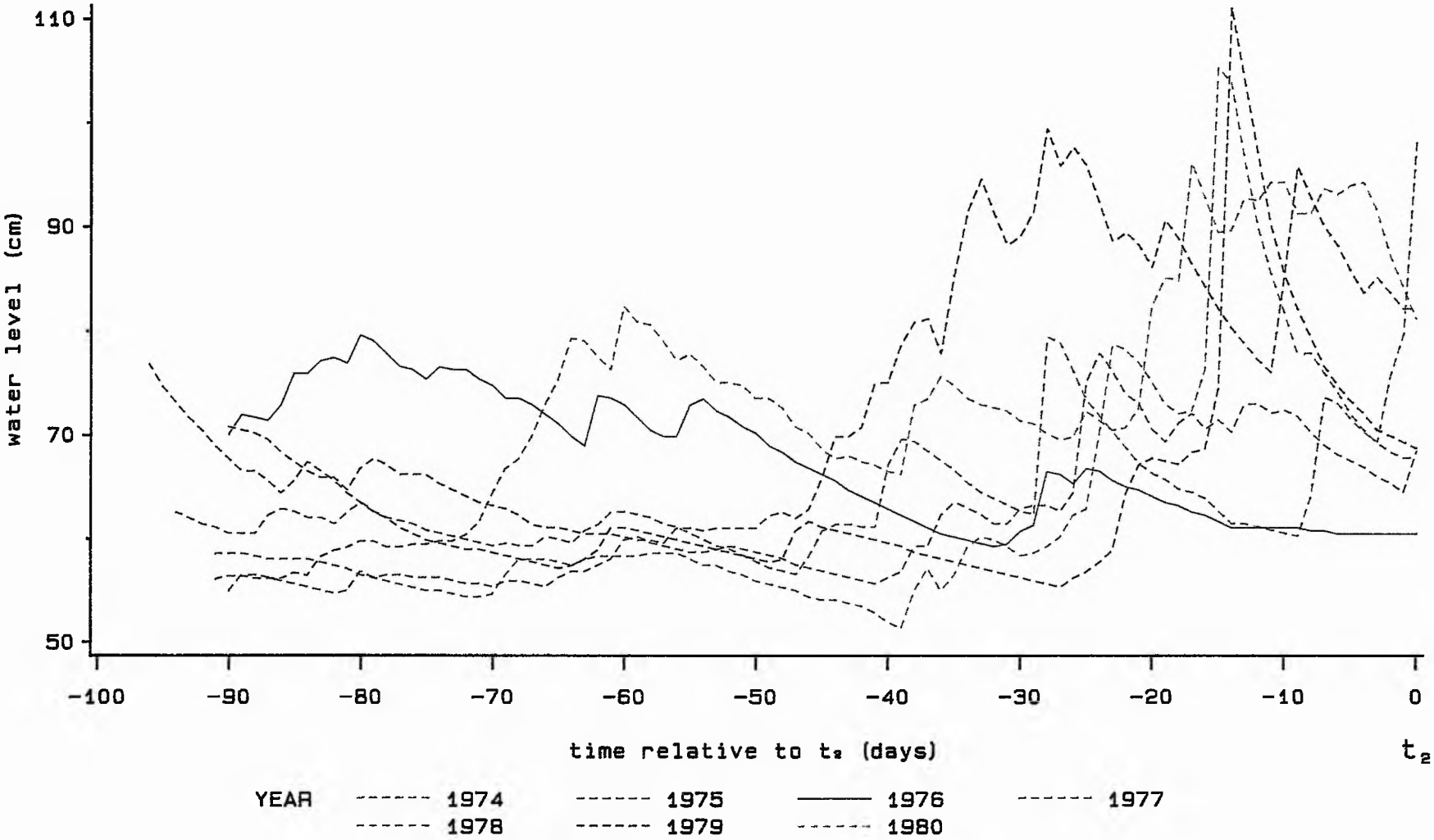


Fig. 6.5b Water level vs. time relative to t_2 1974 - 1980.

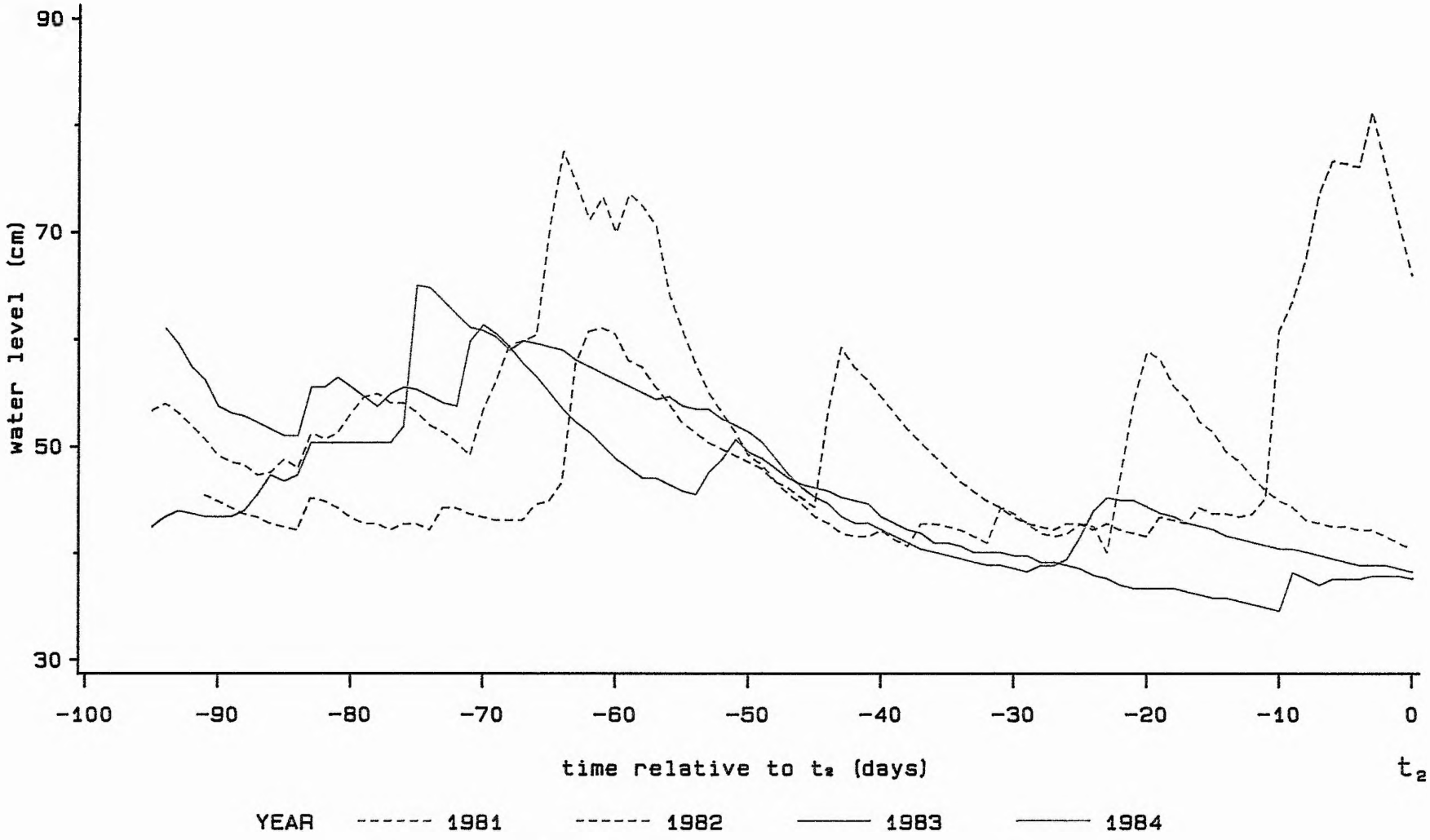


Fig. 6.5c Water level vs. time relative to t_2 1981 - 1984.

water level is low during a "critical" stage of the life cycle; ie during the 30 days before t_2. A possible explanation is that the trout are much larger in the 30 days before t_2 than earlier in the summer, so are affected more by periods of low water level.

6.2.4 Water Level Range

A useful measure of drought based on the Esthwaite water level data must satisfy two criteria:

i) it must measure the severity of a drought; in particular, it must indicate if the water level is low throughout the 30 days before t_2 or if the water level is high at some point in the 30 days before t_2,

ii) it must account for the different mean water levels in the three periods 1967-73, 1974-80, 1981-84.

One possible measure of drought is the mean water level in the 30 days before t_2, adjusted by the mean water level in the relevant time period 1967-73, 1974-80 or 1981-84. However, the estimation of the mean water levels in the time periods 1967-73, 1974-80 and 1981-84 is a non-trivial problem in time series analysis, so this measure is not considered further.

A more convenient measure of drought is based on the range of the water levels in the 30 days before t_2 and is given by w_i = log(maximum water level in the 30 days before t_2 in year i - minimum water level in the 30 days before t_2 in year i), i = 67...84. The measure w satisfies both the criteria above.

First, w is indeed a measure of drought. A study of the water level and rainfall figures (eg those of 1971, Fig. 6.6) shows that the water level rises very quickly after even moderate rainfall. Also, in a period of little rain, the water level falls; the water level falls quickly when the water level is high and very slowly when the water level is low. In a non-drought year, the water level fluctuates greatly, since there is at least one period of high water level in the 30 days before t_2; if there is no rain, the water level falls quickly because the water level is high; if only a moderate amount of rain falls, the water level rises quickly (Fig. 6.7). Hence, the range of the water level is large and w is high. In a drought year, the water level is low throughout the 30 days before t_2; it rains very little, so the water level never rises by much; also, when there is no rain, the water level only drops slightly because the water level is low (Fig. 6.7). Hence, the range of the water level is small and w is low. A graph of w against the mean water level in the 30 days before t_2 (for each of the periods 1967-73, 1974-80, 1981-84) also shows that w is low when the mean water level is low and high when the mean water level is high (Fig. 6.8). Thus, w is a measure of drought, with w decreasing as the severity of a drought increases. (Strictly, w could be low in a non-drought year, since the water level could be high and approximately constant throughout the 30 days before t_2. However, the probability of obtaining quantities of rain that hold the water level high and approximately constant over a 30 day period is very small, since water levels tend to fluctuate greatly when the water level is high. Hence, the situation can effectively be ignored).

Secondly, w accounts for the different mean water levels in the periods 1967-73, 1974-80 and 1981-84. The range of the water levels between 1 May and 30 September does not vary between the three time periods (one way analysis of variance: F =

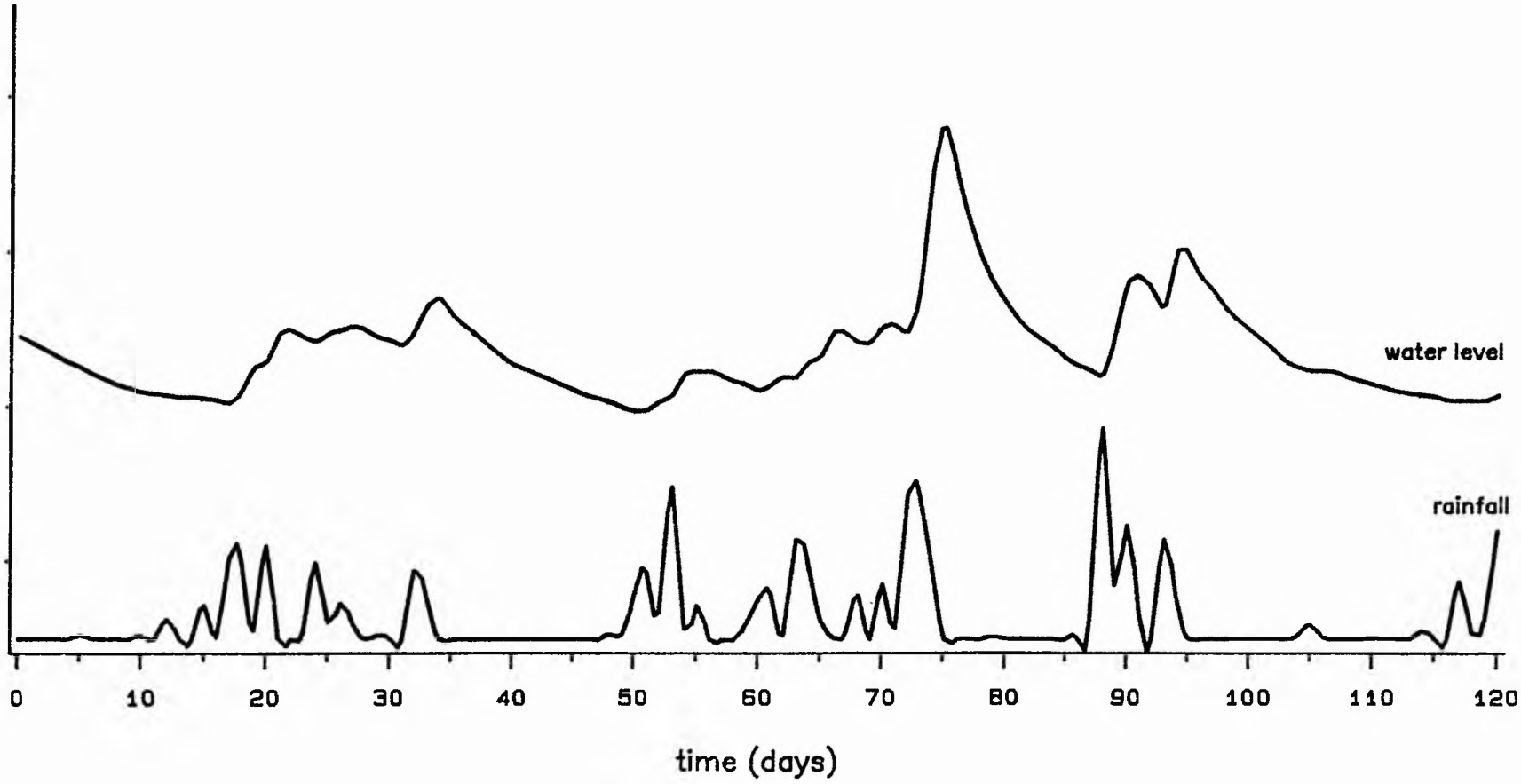


Fig. 6.6 Rainfall and water level in 1971.

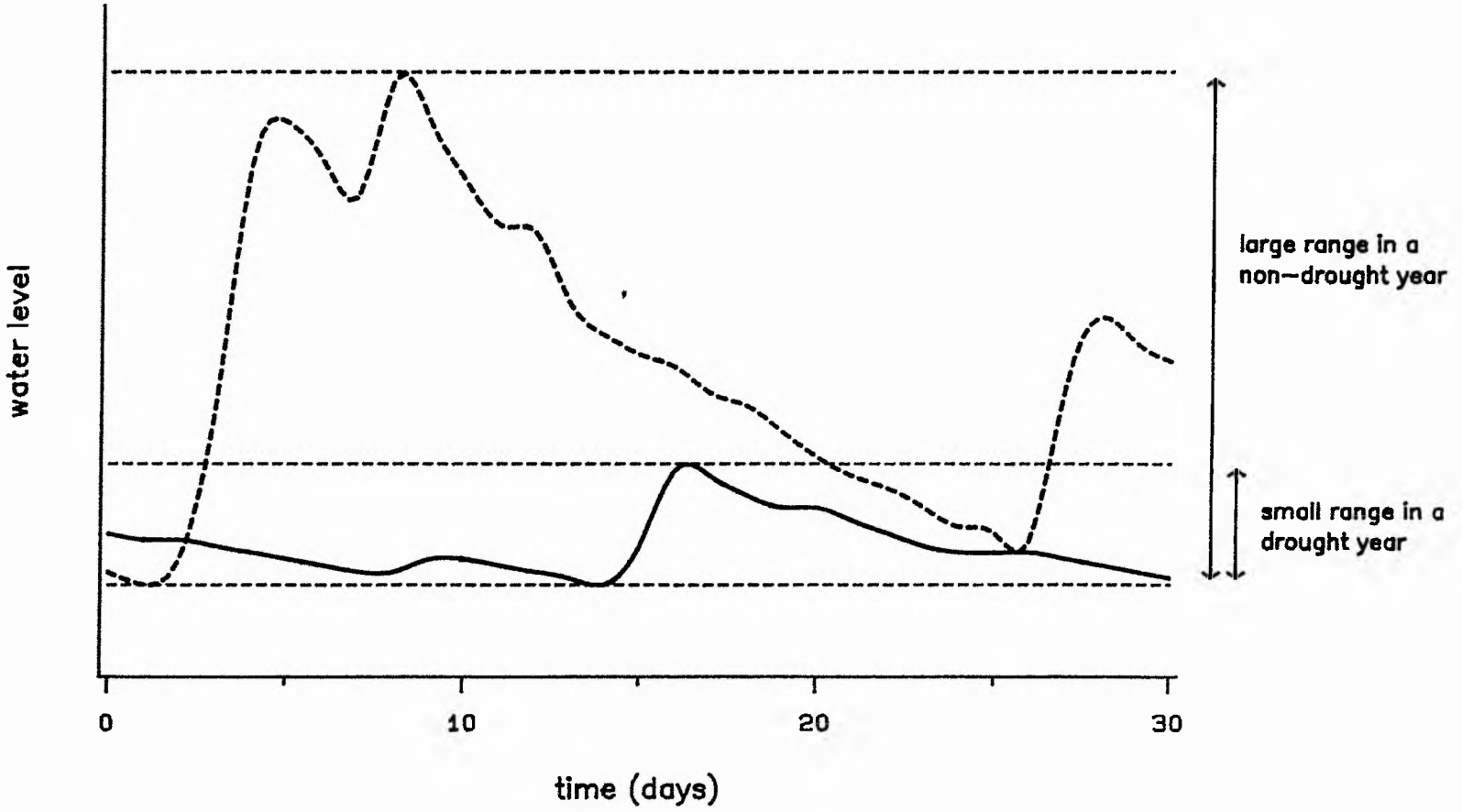


Fig. 6.7 The water level range in a drought and a non—drought year.

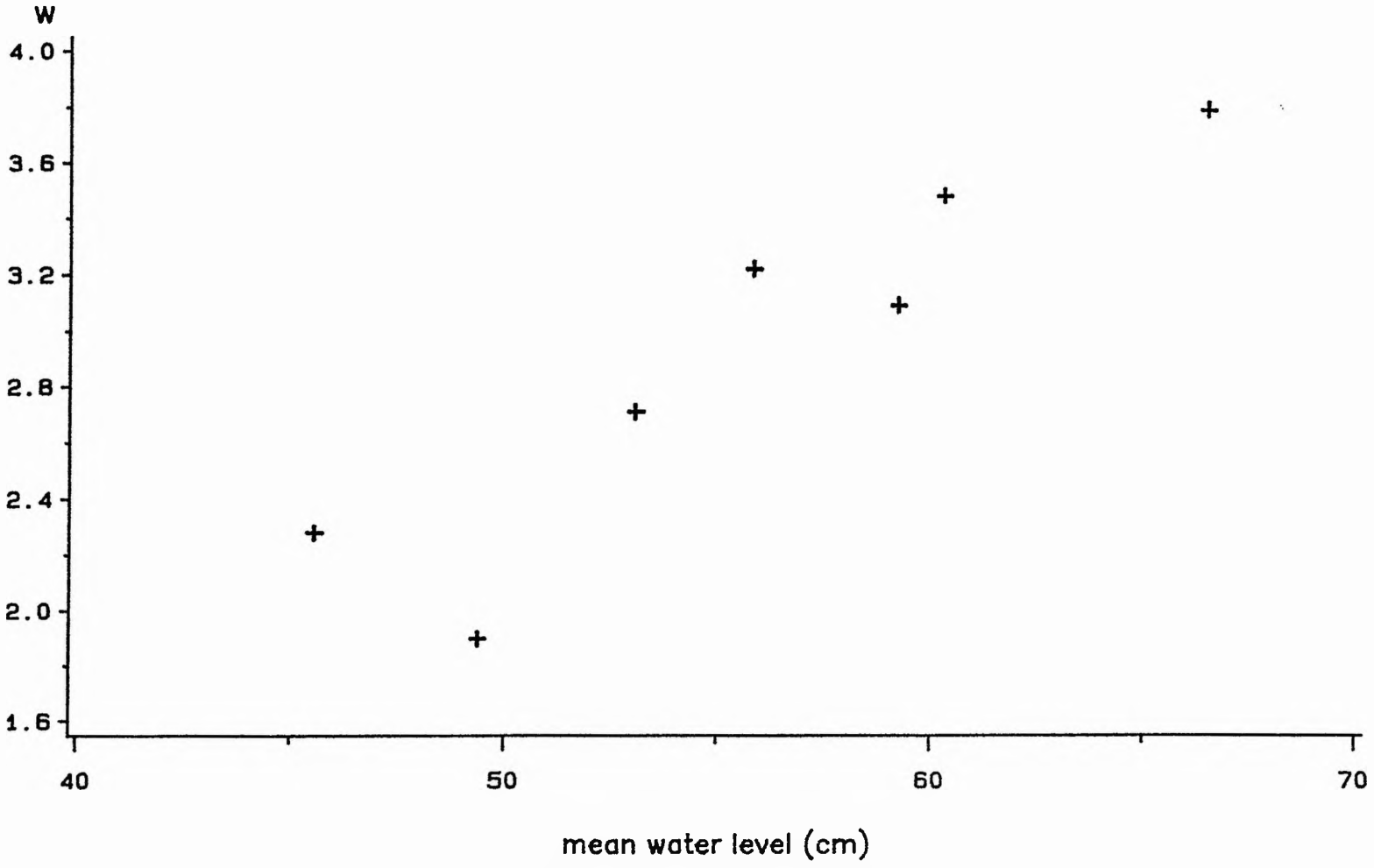


Fig. 6.8a A graph of w against mean water level in the 30 days before t_2 1967 – 1973.

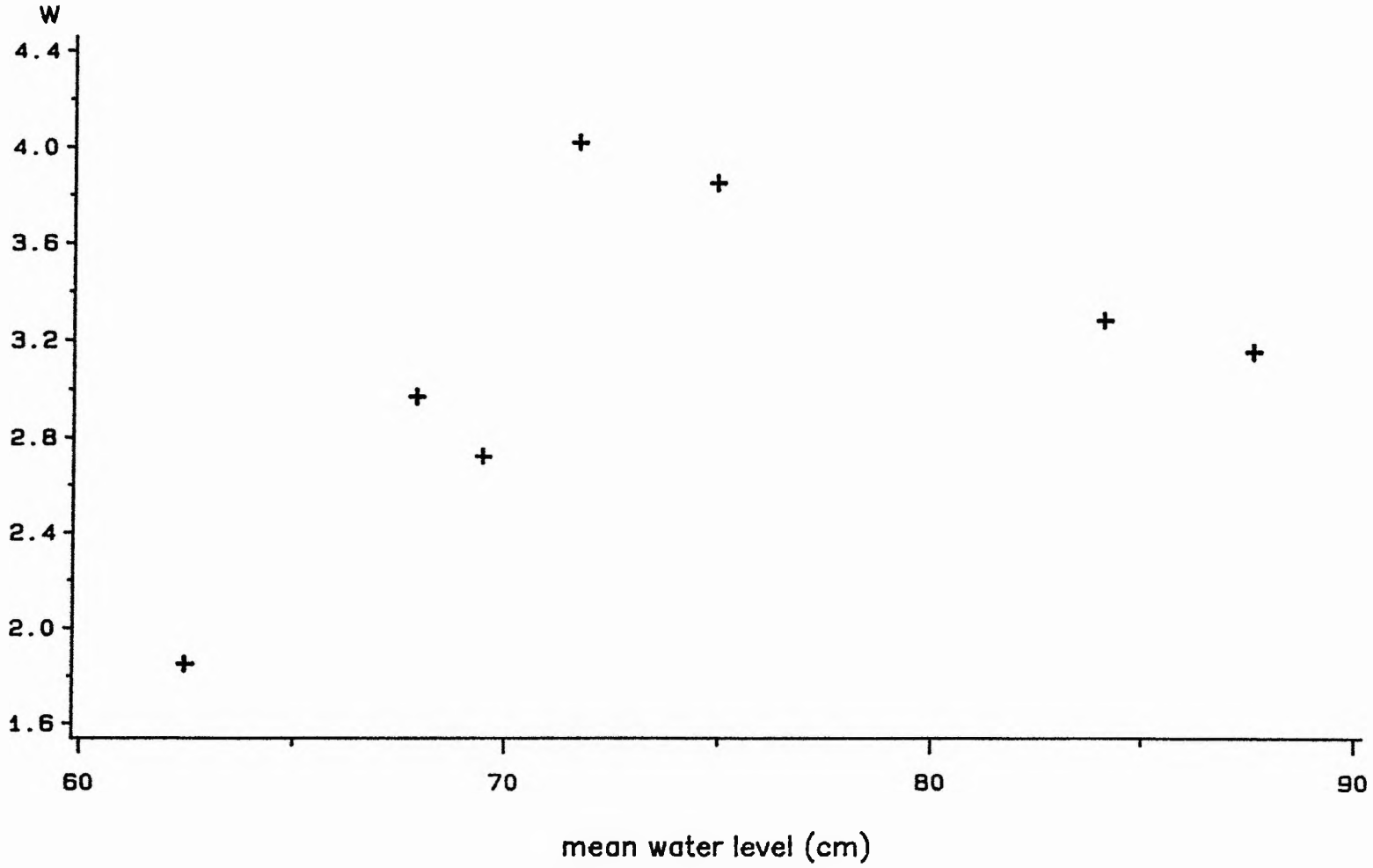


Fig. 6.8b A graph of w against mean water level in the 30 days before t_2 1974 – 1980.

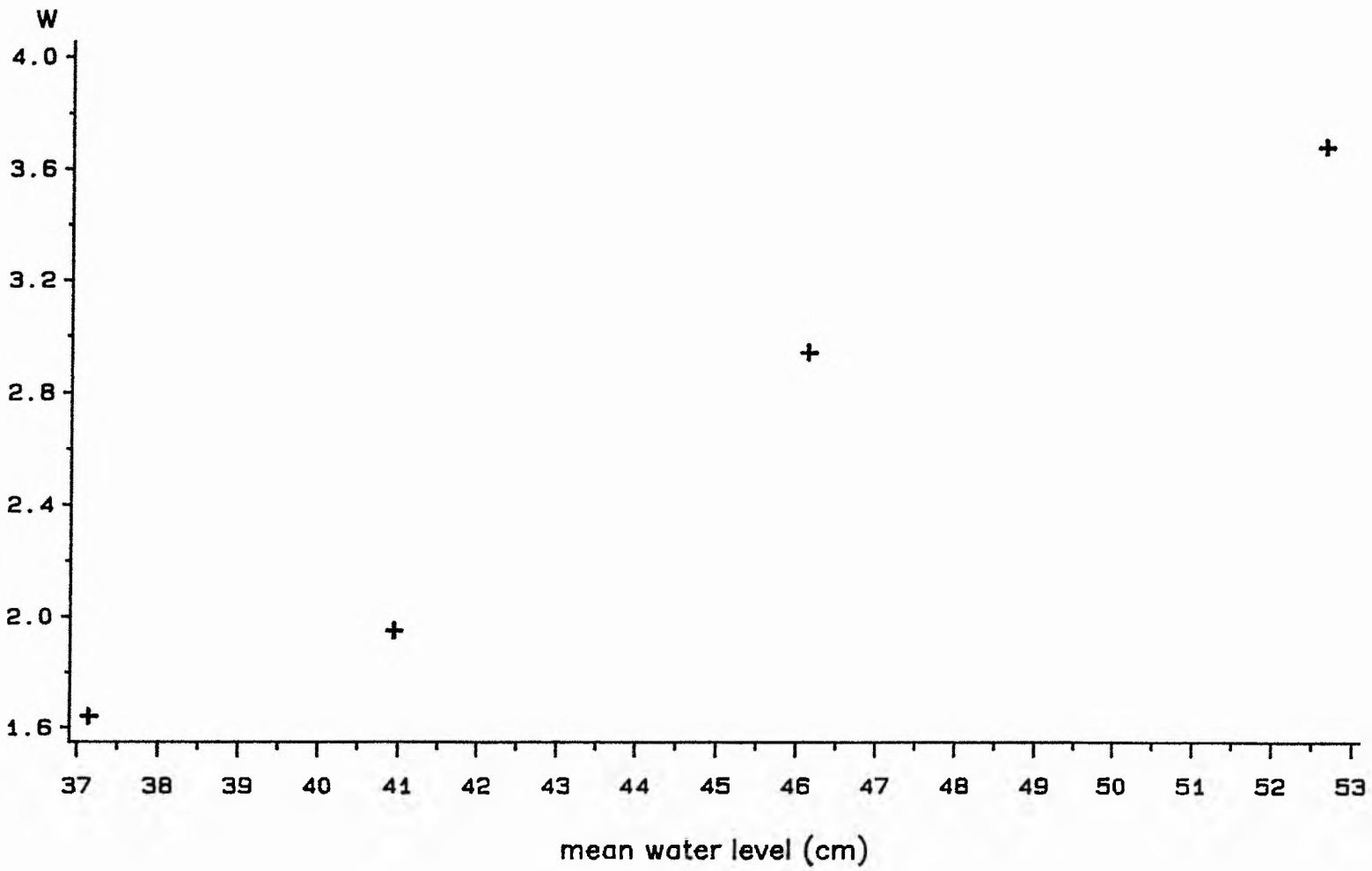


Fig. 6.8c A graph of w against mean water level in the 30 days before t_2 1981 – 1984.

0.34, $F(2,15;0.05) = 3.68$, $p > 0.05$). An examination of the residuals gives no evidence that the assumptions behind the ANOVA are not valid. Thus, it is assumed that the two periods of drainage work do not affect the "behaviour" of the water level of Esthwaite, other than by changing the mean level of the lake. Since w is independent of the mean water level, it can be used to compare droughts across all three time periods.

The log transformation of the water level ranges is taken to reduce the large variation in ranges in the non-drought years. The values $\{w_i\}$ are given in Table 6.2.

6.2.5 Relationship Between R_2, E and W

If there were no droughts in any of the years 1967-84, the expected numbers of trout at t_2 would be estimated to be

$$E_i \exp(-2.400 - 2.790*10^{-4} E_i), \quad i = 67...84$$

(using equation (6.1) with the parameter estimates in Table 6.1). Thus, a measure of the "drought effect" each year (ie the effect of drought on survival) is given by the ratio of the observed number of trout at t_2 to the expected number of trout (in the absence of drought) at t_2,

$$r_i = R_{2i}/\{E_i \exp(-2.400 - 2.790*10^{-4} E_i)\} \quad i = 67...84 \qquad (6.2)$$

In a non-drought year, the observed number of trout is similar to the expected number of trout and r is close to 1. In a drought year, the observed number of

TABLE 6.2

Values of $\{w_i\}$ and $\{r_i\}$

Year	w	r
1967	3.79	1.01
1968	2.28	0.84
1969	1.90	0.76
1970	3.09	1.01
1971	2.71	0.95
1972	3.22	0.96
1973	3.48	1.00
1974	2.72	0.94
1975	4.02	1.02
1976	1.85	0.77
1977	2.97	1.00
1978	3.85	1.00
1979	3.15	1.04
1980	3.28	1.14
1981	2.94	1.06
1982	3.68	1.03
1983	1.64	0.38
1984	1.95	0.14

trout is less than the expected number of trout and r is less than 1. As the effect of drought on survival increases, r decreases towards 0. The values $\{r_i\}$ are given in Table 6.2.

The drought effects $\{r_i\}$ are plotted against the measures of drought severity $\{w_i\}$ in Figure 6.9. In a severe drought, w is low and many fewer trout survive than expected, so r is low. When there is no drought, w is large and r is close to 1. A family of curves which describes such a relationship is given by

$$r = \exp(-\zeta/(w - \xi)), \tag{6.3}$$

since when w is small (strictly, when w is close to but greater than ξ), r is small and when $w \to \infty$, $r \to 1$. For example, the curve

$$r = \exp(-0.09/(w - 1.54))$$

lies close to all the points (w_i,r_i) except that for 1984 (Fig. 6.9). The parameter ξ is interpreted as the severity of a drought in which the survival rate is zero, since $r \downarrow 0$ as $w \downarrow \xi$. (The case $w \leq \xi$ is discussed in Section 6.3.1). The parameter ζ is loosely interpreted as a measure of the rate at which $r \to 1$ as w increases (ie the rate at which the effect of drought on survival decreases as the severity of a drought decreases).

Combining equations (6.2) and (6.3) suggests that the relationship between R_2, E and w might be well described by

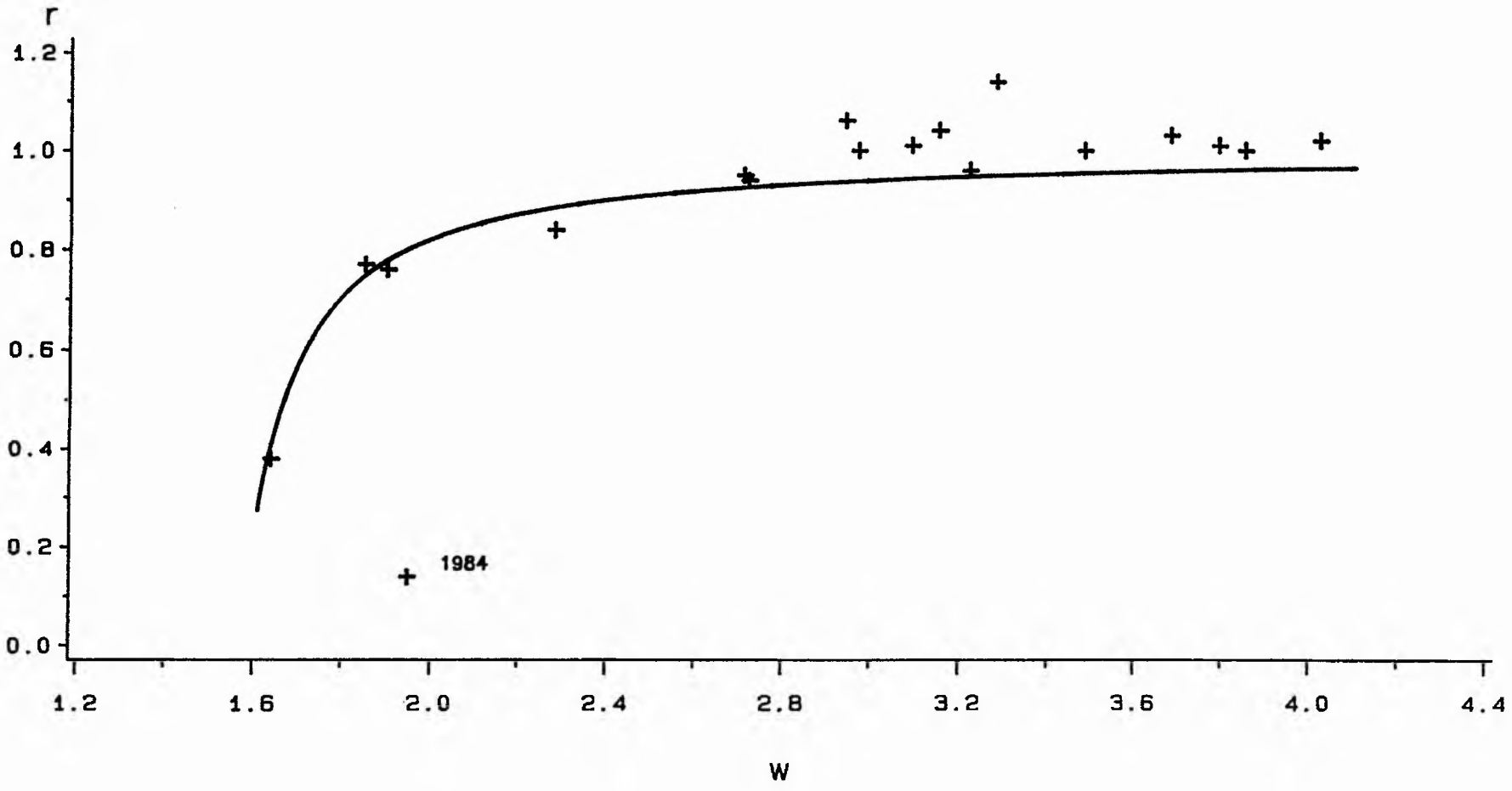


Fig. 6.9 A graph of drought effect r against drought severity w.

$$R_{2i} = E_i \exp(\alpha_2 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_i, \quad i = 67...84 \qquad (6.4)$$

where $\{\epsilon_i\}$ are NID$(0,\sigma^2)$. If model (6.4) is fitted by nonlinear least squares, the 1984 residual is still large and negative. However, if the 1984 data point is removed, model (6.4) gives an excellent fit to the data (Fig. 6.10). An examination of the residuals, including a plot of the residuals against $\{w_i\}$, gives no evidence that the error structure is not appropriate and reveals no systematic departures of the data from the model. Parameter estimates are given in Table 6.3. The maximum intrinsic curvature $\Gamma^N = 0.11$ is less than $1/(2\sqrt{F(4,13;0.05)}) = 0.28$, so the tangent plane approximation to the solution locus is acceptable in the 95% confidence region of interest. Hence, the residual sum of squares obtained by fitting (6.4) with

i) the parameters unconstrained

ii) the constraint $\zeta_2 = 0$

can be compared to show that the inclusion of the "drought term" significantly improves the fit of the model ($F = 273.7$, $F(1,13;0.001) = 17.81$, $p < 0.001$).

A further investigation of the water level figures and a comparison of the Ambleside rainfall figures (and also temperature figures) for 1984 with those of other years, revealed no explanation for the extremely low survival rate in 1984, except, perhaps, that it was due to an unusually dry early spring. Hence, the 1984 data point is omitted throughout the rest of this Chapter.

The maximum parameter effects curvature of model (6.4) (with the 1984 data point omitted) is large compared to $1/(2\sqrt{F})$ and the RMS parameter effects curvature is

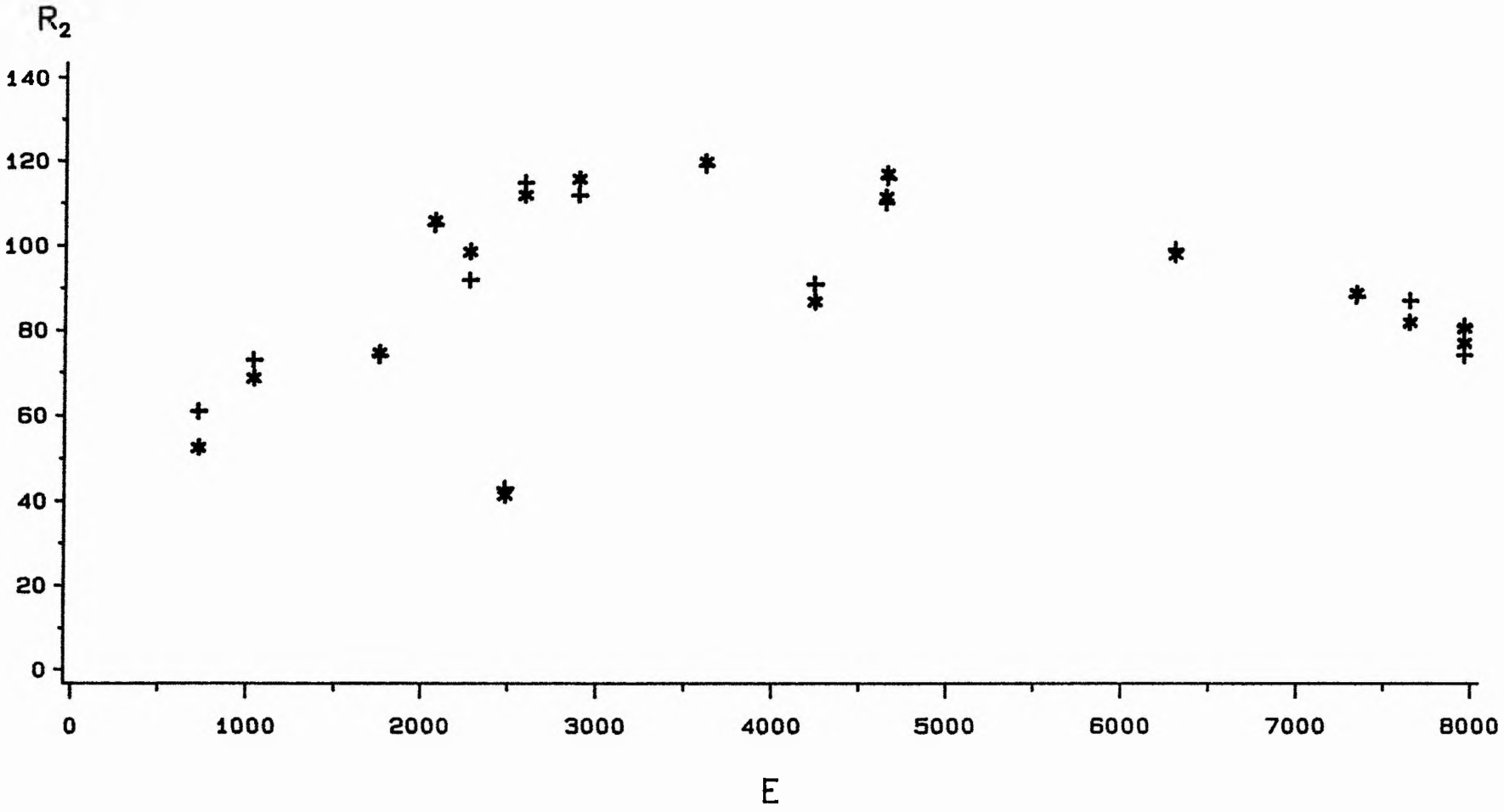


Fig. 6.10 Observed numbers of 0+ parr at t_2 (+) and fitted values from model (6.4) (*).

TABLE 6.3

Parameter estimates for model (6.4)

Parameter	Estimate	Lin approx s.e.
α_2	-2.352	0.032
β_2	$2.747*10^{-4}$	$0.051*10^{-4}$
ζ_2	0.1206	0.0255
ξ_2	1.526	0.028

rss = 213.25 on 13 d.f.
s^2 = 16.40

close to $1/(2\sqrt{F})$ (Table 6.4) so the uniform co-ordinates approximation is not acceptable in the 95% region of interest. However, the parameter effects nonlinearity is greatly reduced by a pair of "expected value" transformations

$$\gamma_{\omega j} = \exp(-\zeta_2/(\omega_j - \xi_2)), \qquad j = 1,2,$$

for specified values of ω_1, ω_2, where $\gamma_{\omega 1}$ and $\gamma_{\omega 2}$ are interpreted as the "expected" drought effects when $w = \omega_1$ and $w = \omega_2$ respectively. Suitable values of ω_1 and ω_2 (found by trial and error) are $\omega_1 = 1.64$ and $\omega_2 = 2.28$; these are the values of w in 1983 and 1968, so $\gamma_{\omega 1}$ and $\gamma_{\omega 2}$ are interpreted as the expected drought effects in a "serious" drought and a "moderately serious" drought respectively. The drought model is written in terms of α_2, β_2, $\gamma_{\omega 1}$ and $\gamma_{\omega 2}$ as

$$R_{2i} = E_i \exp\left(\alpha_2 - \beta_2 E_i + \frac{(\omega_2 - \omega_1)\log(\gamma_{\omega 1})\log(\gamma_{\omega 2})}{(w_i - \omega_1)\log(\gamma_{\omega 1}) - (w_i - \omega_2)\log(\gamma_{\omega 2})}\right) + \epsilon_i$$

$$i = 67...83 \qquad (6.5)$$

The maximum parameter effects curvature for this parameterisation is less than $1/(2\sqrt{F})$ (Table 6.4) so the uniform co-ordinates approximation is acceptable in the 95% region of interest.

Nonlinearity measures for the parameters α_2, β_2, ζ_2, ξ_2, $\gamma_{\omega 1}$ and $\gamma_{\omega 2}$ are given in Table 6.5. The behaviour of $\hat{\zeta}_2$ and $\hat{\xi}_2$ is very nonlinear. However, the behaviour of $\hat{\alpha}_2$, $\hat{\beta}_2$, $\hat{\gamma}_{\omega 1}$ and $\hat{\gamma}_{\omega 2}$ is much closer-to-linear, although the distribution of $\hat{\gamma}_{\omega 2}$ is still skewed. A number of monotonic transformations of $\gamma_{\omega 2}$ were investigated to reduce this asymmetry, but all considerably increased the parameter

TABLE 6.4

Curvature measures for models (6.4) and (6.5)

Maximum intrinsic curvature = 0.11		
RMS intrinsic curvature = 0.04		
Axis ratios: likelihood	1.00	1.04
95% confidence	1.00	1.12
Quadratic approximation	0.93%	
	Parameterisation (6.4)	(6.5)
Maximum parameter effects curvature	0.57	0.26
RMS parameter effects curvature	0.24	0.09
$1/(2\sqrt{F(4,13;0.05)}) = 0.28$		

TABLE 6.5

Asymmetry measures of nonlinearity for models (6.4) and (6.5)

Parameter	α_2	β_2
Least squares estimate	-2.352	$2.747*10^{-4}$
Lin approx s.e.	0.032	$0.051*10^{-4}$
Box's bias	0.001	$2*10^{-8}$
Hougaard's skewness	0.049	0.013
Lowry and Morton's λ	0.005	$3*10^{-4}$
Simulation measures:		
True value	-2.352	$2.747*10^{-4}$
Variance of l.s.e. (lin approx)	$1.014*10^{-3}$	$2.636*10^{-11}$
Simulated value	-2.352	$2.746*10^{-4}$
Simulated variance	$1.076*10^{-3}$	$2.625*10^{-11}$
T_3 - skewness stat	1.18	1.46
T_4 - excess kurtosis stat	0.44	-0.62
Lowry and Morton's λ	0.008	0.003

Parameter	ζ_2	ξ_2
Least squares estimate	0.1206	1.526
Lin. approx. s.e.	0.0255	0.028
Box's bias	0.0026	-0.003
Hougaard's skewness	0.588 **	-0.742 **
Lowry and Morton's λ	0.027 *	0.036 *
Simulation:		
True value	0.1206	1.526
Variance of l.s.e. (lin approx)	$6.511*10^{-4}$	$7.911*10^{-4}$
Simulated value	0.1225 *	1.523 **
Simulated variance	$7.728*10^{-4}$ **	$9.393*10^{-4}$ **
T_3 - skewness stat	9.48 **	-10.90 **
T_4 - excess kurtosis stat	6.81 **	7.80 **
Lowry and Morton's λ	0.030 *	0.038 *

TABLE 6.5 (continued)

Parameter	$\gamma_{\omega 1}$	$\gamma_{\omega 2}$
Least squares estimate	0.3638	0.8518
Lin approx s.e.	0.0337	0.0243
Box's bias	-0.0005	-0.0013
Hougaard's skewness	0.008	-0.279 *
Lowry and Morton's λ	0.002	0.010
Simulation:		
True value	0.3638	0.8518
Variance of l.s.e. (lin approx)	$1.134*10^{-3}$	$5.895*10^{-4}$
Simulated value	0.3640	0.8514
Simulated variance	$1.136*10^{-3}$	$6.521*10^{-4}$ *
T_3 - skewness stat	-2.24 *	-4.69 **
T_4 - excess kurtosis stat	0.58	1.83
Lowry and Morton's λ	0.004	0.012 *
95% confidence intervals:		
Lin approx	0.298 0.430	0.804 0.899
Simulation	0.297 0.428	0.799 0.896

* indicates moderate nonlinear behaviour
** indicates serious nonlinear behaviour

effects curvatures and so were not considered further. Since the parameter effects curvatures for model (6.5) are less than $1/(2\sqrt{F})$ and since a simulation 95% confidence interval for $\Upsilon_{\omega 2}$ is very similar to a linear approximation 95% confidence interval (Table 6.5), it is unlikely that the asymmetry of $\hat{\Upsilon}_{\omega 2}$ indicates serious nonlinear behaviour in the 95% region of interest. Thus, it is concluded that model (6.5) is a good model for describing the relationship between R_2, E and w. Parameter estimates are given in Table 6.6. (Strictly, the simulation and linear approximation confidence limits can not be compared, since in the simulation, the variance σ^2 is assumed known a priori, whereas in the linear approximation, uncertainty about the true value of σ^2 is incorporated by using a t distribution to calculate the confidence limits. To compensate for this, the "informal" comparison above uses a linear approximation confidence interval given by $\hat{\Upsilon}_{\omega 2} \pm 1.96$ s.e. rather than the usual $\hat{\Upsilon}_{\omega 2} \pm t(\nu;0.05)$ s.e.).

It is noted that the 1983 data point is very influential in the estimation of $\Upsilon_{\omega 1}$ and $\Upsilon_{\omega 2}$ (or of ξ_2 and ξ_2). However, if the 1983 point is omitted, model (6.5) is still a significantly better fit to the R_2 data than the original Ricker model ($F = 69.3$, $F(1,12;0.001) = 18.6$, $p < 0.001$) and the parameter estimates are virtually unchanged (although the standard errors have inevitably increased). Thus, it is unlikely that the form of (6.5) is spuriously determined by the 1983 data point.

A number of other models relating R_2 to E and w were also investigated. These included models that incorporated logit and complementary log-log functions in both w and $\exp(w)$. However, no model was found that gave as good a fit to the R_2 data as model (6.5). The difficulty arises in accounting for the gradual decrease in the survival rate as w decreases from about 2.5 to 1.8 (ie from no drought to a

TABLE 6.6

Parameter estimates for model (6.5)

Parameter	Estimate	Lin approx s.e	Lin approx correlations			
α_2	-2.352	0.032	1.00			
β_2	$2.747*10^{-4}$	$0.051*10^{-4}$	0.76	1.00		
$\gamma_{\omega 1}$	0.3638	0.0337	-0.13	-0.12	1.00	
$\gamma_{\omega 2}$	0.8518	0.0243	-0.66	-0.16	0.02	1.00

rss = 213.25 on 13 d.f.
s^2 = 16.40

moderately serious drought), followed by the sharp decrease in the survival rate as w decreases from about 1.8 to 1.6 (ie from a moderately serious drought to a serious drought) (cf. Fig. 6.9).

6.2.6 Rainfall

The total rainfall in the i days before t_2, i = 1...60, is shown in Figure 6.11 for the years 1967-84. The total rainfall in the 60 days before t_2 is low in all the drought years and high in most of the non-drought years. This is consistent with the results from the water level data, since if the water level is low in the 30 days before t_2, the rainfall in the 60 days before t_2 must also be low. However, the converse is not true. For example, the total rainfall in the 60 days before t_2 in the non-drought year 1981 was similar to that in the drought years 1968 and 1969. However, the distribution of the rainfall over the 60 days was very different. In 1981, most of the rain fell in one spell about 25 days before t_2, so that the water level was quite high about 20 days before t_2. However, in 1968 and 1969 the rainfall was more evenly distributed, so that the water level was low throughout the 30 days before t_2. It is dangerous to draw too many conclusions from single observations such as this. However, it would appear that the short period of high water level in the 30 days before t_2 in 1981 was sufficient to prevent mortality due to the low total rainfall. Thus, both the total rainfall and the distribution of rainfall over the 60 days before t_2 are important factors affecting survival. It was not possible to develop a simple measure of rainfall that accommodated both these features and that could be used to give a satisfactory description of survival to t_2.

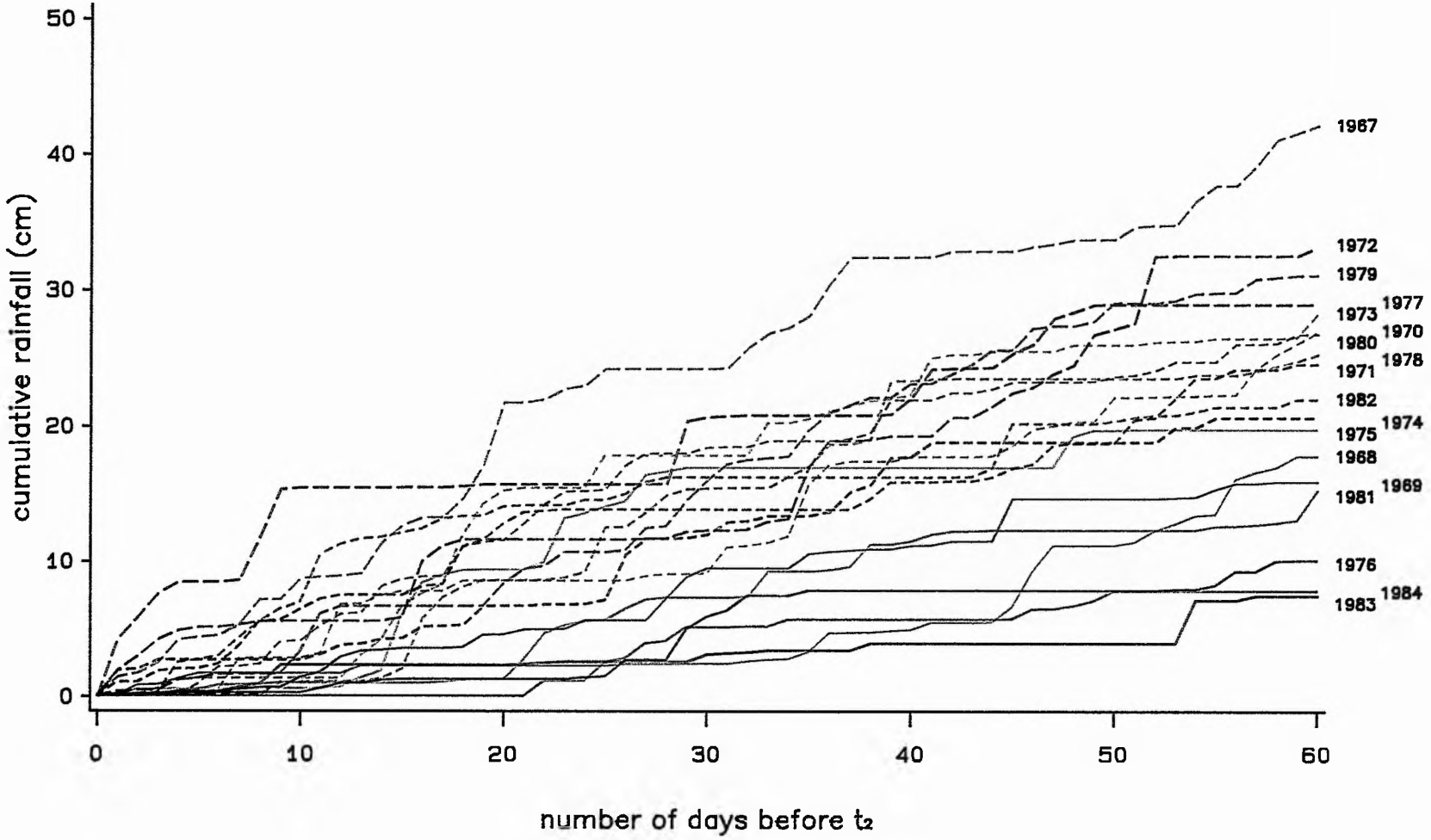


Fig. 6.11 Cumulative rainfall in the days before t_2.

6.2.7 Density Dependent Survival Between t_1 and t_2

The relationships between R_1 and E and between R_2 and E are well described by the models

$$R_{1i} = E_i \exp(\alpha_1 - \beta_1 E_i) + \epsilon_{1i}, \quad i = 67...84, \tag{6.6}$$

where (ϵ_{1i}) are NID$(0,\sigma_1^2)$, and

$$R_{2i} = E_i \exp(\alpha_2 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{2i}, \quad i = 67...83 \tag{6.7}$$

where (ϵ_{2i}) are NID$(0,\sigma_2^2)$. (The parameters ζ_2 and ξ_2 are used in preference to $\Upsilon_{\omega 1}$ and $\Upsilon_{\omega 2}$ since the notation is less cumbersome and since the arguments that follow do not depend on the parameterisation of the "drought" term in (6.7)).

The parameters β_1 and β_2 are measures of density dependent survival from the time the eggs are laid to t_1 and t_2 respectively (Section 2.3). However, since there is negligible mortality between the time the eggs are laid and t_0, the time of the alevins sample (Elliott, 1984a), β_1 and β_2 are equivalent to measures of density dependent survival from t_0 to t_1 and from t_0 to t_2. The value of $\beta_2 - \beta_1$ is of particular biological interest. If $\beta_1 = \beta_2$, E can be eliminated from (6.6) and (6.7) to give

$$R_{2i} = R_{1i} \exp(\alpha_2 - \alpha_1 - \zeta_2/(w_i - \xi_2)) + \epsilon_{12i},$$

where

$$\epsilon_{12i} = \epsilon_{2i} - \epsilon_{1i}\exp(\alpha_2 - \alpha_1 - \zeta_2/(w_i - \xi_2)), \quad i = 67...83,$$

in which case, density dependent survival occurs only between t_0 and t_1 and there is proportional survival between t_1 and t_2. However, if $\beta_1 \neq \beta_2$, R_2 is not proportional to R_1 and there is density dependent survival between t_1 and t_2. (In this case, the relationship between R_1 and R_2 is of the form

$$\log(R_{2i}/R_{1i}) = \alpha_2 - \alpha_1 - \zeta_2/(w_i - \xi_2) - K\left(R_{1i}^{\beta_2} R_{2i}^{-\beta_1}\right)^{1/(\beta_2-\beta_1)}, \qquad (6.8)$$

where

$$K = (\beta_2 - \beta_1)\exp\left(\frac{1}{\beta_2 - \beta_1}\ (\alpha_2\beta_1 - \alpha_1\beta_2 - \beta_1\zeta_2/(w_i - \xi_2))\right)$$

and where terms involving the errors ϵ_{1i} and ϵ_{2i} have been omitted for simplicity).

The values of β_1 and β_2 are compared by combining (6.6) and (6.7) into one model. The two sets of residuals obtained by fitting models (6.6) and (6.7) by nonlinear least squares are positively correlated ($\rho = 0.628$, df = 15, $p < 0.01$). Hence, it is assumed that the joint relationship between R_1, R_2 and E can be described by

$$R_{1i} = E_i\exp(\alpha_1 - \beta_1 E_i) + \epsilon_{1i}, \quad i = 67...84,$$

$$R_{2i} = E_i\exp(\alpha_2 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{2i},$$

$$i = 67...83, \qquad (6.9)$$

with

i) $\varepsilon_{1i}, \varepsilon_{2i}$ independent of $\varepsilon_{1j}, \varepsilon_{2j}$ if $i \neq j$

ii) $$\begin{pmatrix}\epsilon_{1i}\\ \epsilon_{2i}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right), \quad i = 67...83$$

$$\epsilon_{1,84} \sim N(0,\sigma_1^2)$$

The likelihood of (6.9) is maximised with

i) all the parameters unconstrained, $-2\log L = 166.42$,

ii) the constraint $\beta_1 = \beta_2$, $-2\log L = 235.79$.

Thus, $\beta_1 \neq \beta_2$ ($-2\log\lambda = 69.37, \chi^2(1;0.001) = 10.8$, $p < 0.001$), so survival between t_1 and t_2 is density dependent.

The approach described above is more satisfactory than the "key factor analysis" used by Elliott (1985a) to investigate density dependence between t_1 and t_2. Elliott assumes that the loss rate between t_1 and t_2, $\log(R_1/R_2)$, is related to R_1 by

$$\log(R_{1i}/R_{2i}) = a + bR_{1i} + \epsilon_i, \quad i = 67...83 \qquad (6.10)$$

where $\{\varepsilon_i\}$ are NID$(0,\sigma^2)$, and then shows that survival between t_1 and t_2 is density dependent since b is significantly different from zero. It is clear that $\log(R_1/R_2)$ is positively correlated with R_1 ($\rho = 0.794$, df = 15, $p < 0.001$). However, as Elliott admits, a drawback of his analysis is that the initial assumption that $\log(R_1/R_2)$ and R_1 are related through (6.10) is not strictly valid. Rather, R_1 and R_2 are related by a curve of the form (6.8) (cf Fig. 9.1). Also, (6.10) fails to account for the effect of drought on survival. The advantage of the method used in this section is that there is no evidence that the "maximal" model (6.9) is not valid. Model (6.9) is

based on models (6.6) and (6.7) which give excellent descriptions of the relationships between R_1 and E and between R_2 and E (and account for the drought effect). Further, (6.9) allows for correlations between R_1 and R_2 within each year class.

The survival rate between t_1 and t_2 is investigated in more detail in Chapter 9.

6.3 R_3 - 1+ PARR (MAY/JUNE)

6.3.1 Survival Between t_0 and t_3

The water level of Esthwaite rises sharply, shortly after t_2, in all the drought years except 1976 (Fig. 6.12), so the droughts (except that of 1976) effectively finish at t_2. Therefore, it is assumed that the severity of an entire drought (rather than the portion of the drought to t_2) can be measured by w. The model

$$R_{3i} = E_i \exp(\alpha_3 - \beta_3 E_i - \xi_3/(w_i - \xi_3)) + \epsilon_i, \quad i = 67...83, \qquad (6.11)$$

where $\{\epsilon_i\}$ are NID$(0, \sigma^2)$, is fitted by nonlinear least squares and is a good description of the relationship between R_3, E and w. An examination of the residuals gives no evidence that the error structure is not appropriate. A plot of the residuals against w reveals no systematic departures. Even the 1976 data point is well described. Parameter estimates are given in Table 6.7.

The intrinsic nonlinearity of model (6.11) is moderately large, since the maximum intrinsic curvature is greater than $1/(2\sqrt{F})$ and the axis ratios are not close to 1 (Table 6.8). Thus, any inference based on a linear approximation must be treated with caution. In particular, linear approximation 95% confidence regions for both

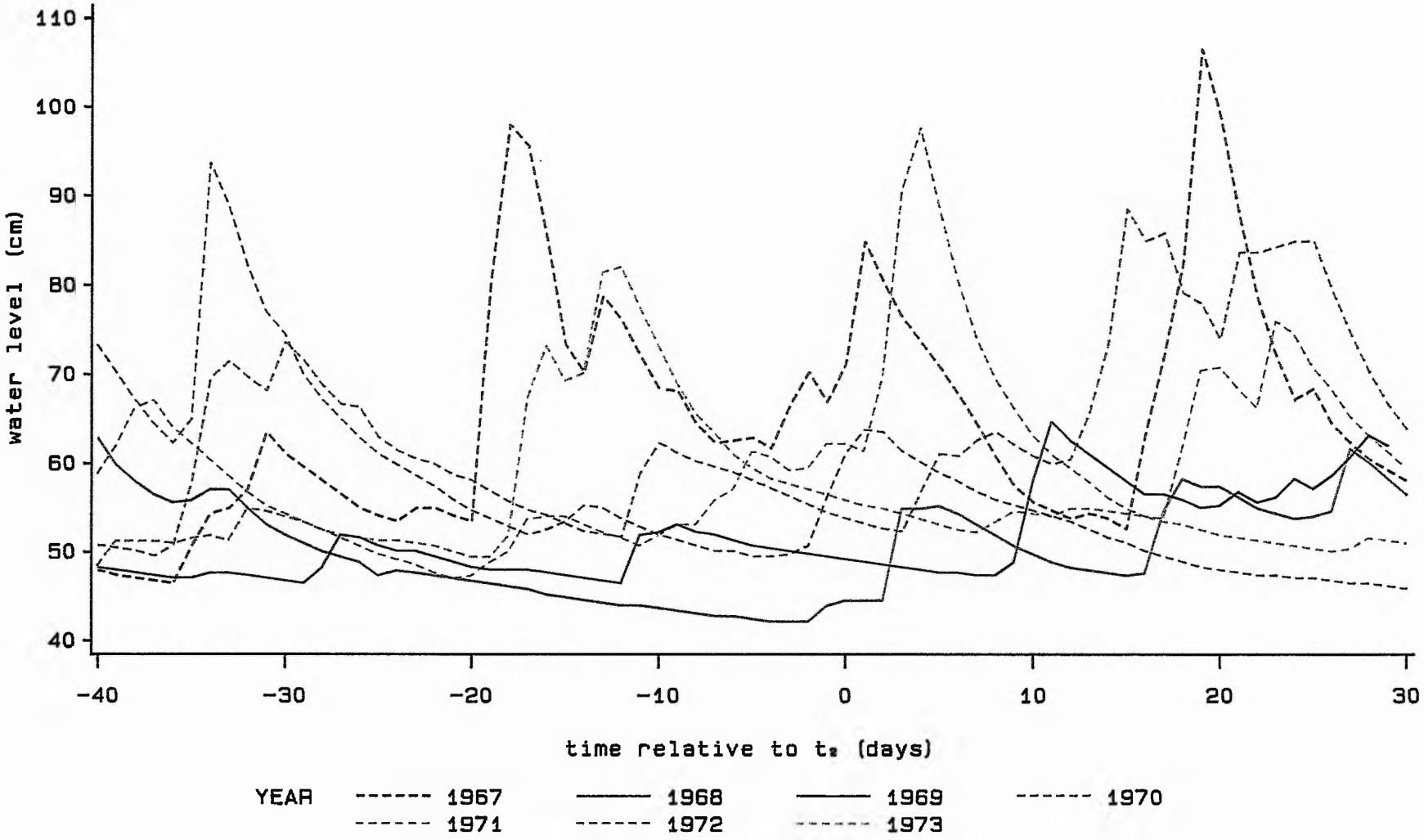


Fig. 6.12a Water level vs. time relative to t_s 1967 - 1973.

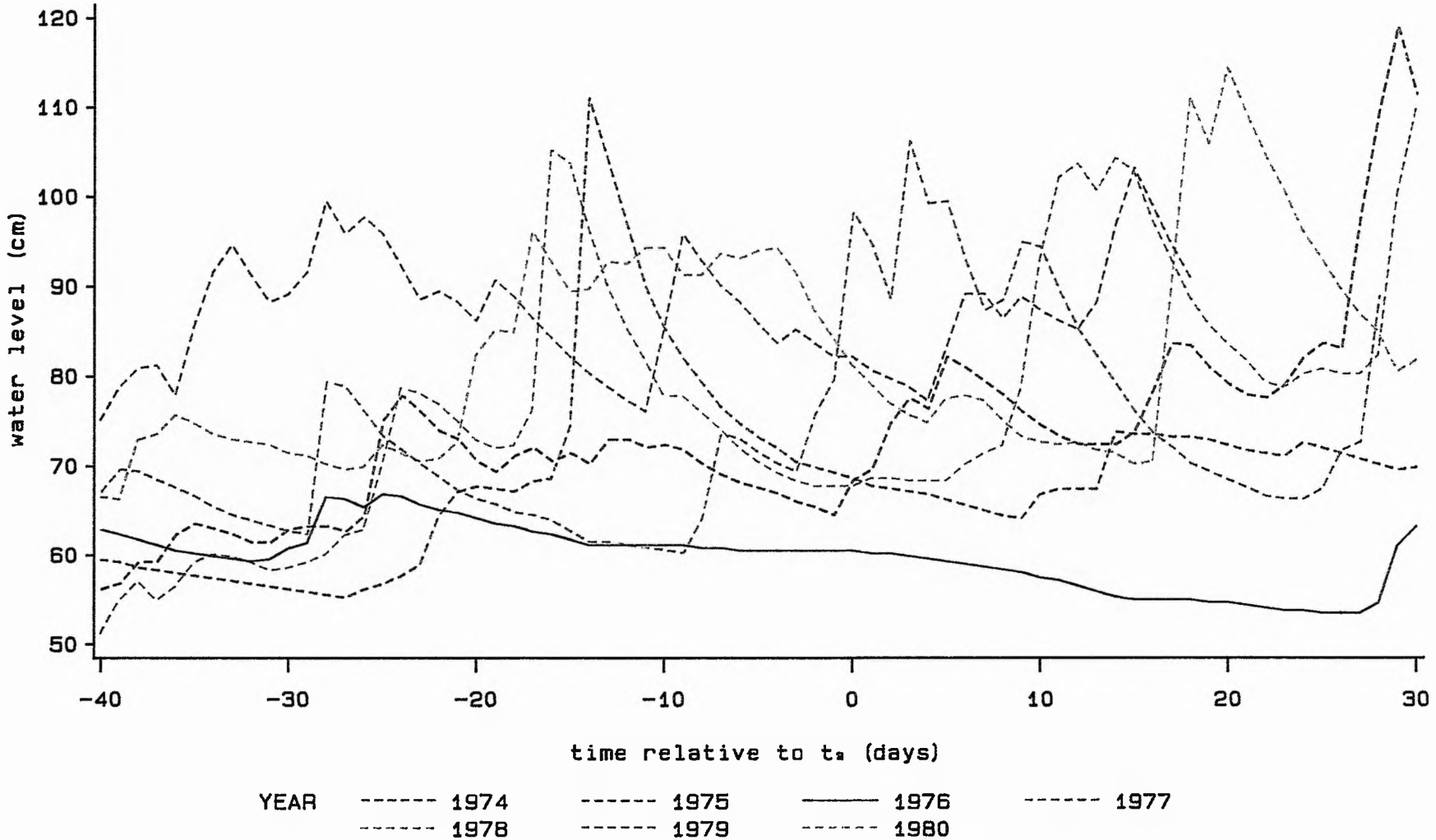


Fig. 6.12b Water level vs. time relative to t_s 1974 - 1980.

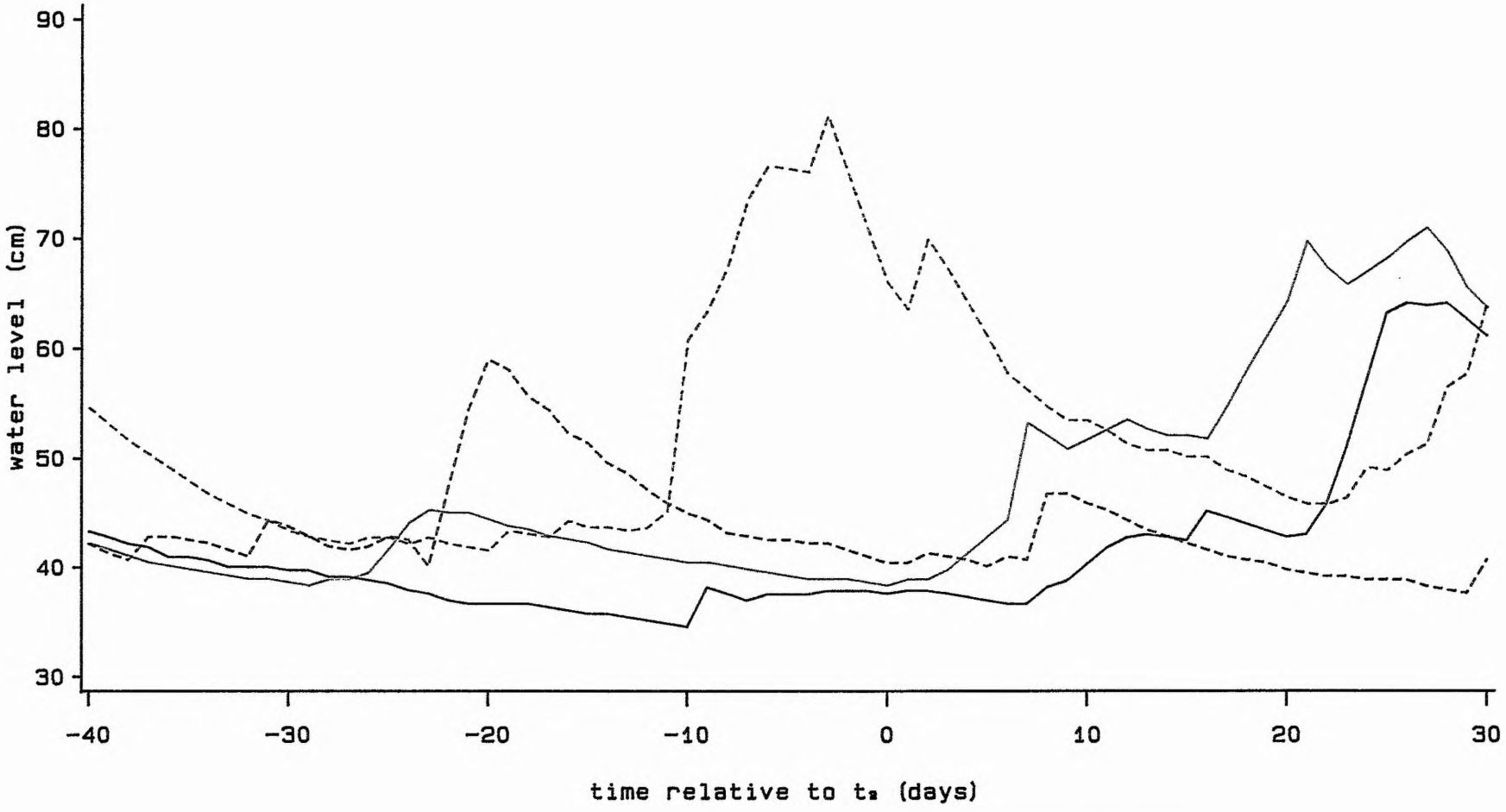


Fig. 6.12c Water level vs. time relative to t_a 1981 - 1984.

TABLE 6.7

Parameter estimates for model (6.11)

Parameter	Estimate	Lin approx s.e.	95% confidence limits	
α_3	-3.270	0.127	-3.511	-2.947
β_3	$2.691*10^{-4}$	$0.217*10^{-4}$	$2.299*10^{-4}$	$3.130*10^{-4}$
ζ_3	0.110	0.078	$7.511*10^{-4}$	0.909
ξ_3	1.566	0.067	0.830	1.849

rss = 592.61 on 13 d.f.
s^2 = 45.59

TABLE 6.8

Curvature measures for model (6.11)

Maximum intrinsic curvature = 0.32		
RMS intrinsic curvature = 0.12		
Axis ratios: likelihood	1.00	1.05
95% confidence	0.99	1.15
Quadratic approximation	7.98%	
	Parameterisation	
	$(\alpha_3, \beta_3, \zeta_3, \xi_3)$	$(\alpha_3, \beta_3, \gamma_{\omega 1}, \gamma_{\omega 2})$
Max parameter effects curvature	1.90	0.76
RMS parameter effects curvature	0.75	0.27
$1/2(\sqrt{F(4,13;0.05)}) = 0.28$		

sets of parameters (α_3, β_3, ζ_3, ξ_3) and (α_3, β_3, $\gamma_{\omega 1}$, $\gamma_{\omega 2}$) are totally inadequate since the relevant parameter effects curvatures are very large (Table 6.8). However, model (6.11) is not well defined (as discussed below) and this problem must be addressed before inferences based on (6.11) can be investigated.

Model (6.11) is not well defined in two ways. First, the model allows negative numbers of trout, which is unsatisfactory from a biological viewpoint. Secondly, the model is not defined at $w = \xi_3$ and makes little sense if $w < \xi_3$. Strictly, model (6.4) is not well defined for the same reasons. In the case of model (6.4), the problem can be ignored, since

i) given the parameter values in Table 6.3, the probability of obtaining negative numbers of trout at t_2 is negligible,

ii) the 99% confidence region of interest does not contain values of $\xi_2 \geq 1.64 = \min\{w_i, i = 67...83\}$; that is, it is extremely unlikely that any of the observed values of w is less than the "true" value of ξ_2.

However, in the case of model (6.11),

i) given the parameter values in Table 6.7, there is a non-negligible probability (about 0.05) of obtaining negative numbers of trout at t_3,

ii) the point $\xi_3 = 1.64$ is contained in the 95% region of interest.

Hence, model (6.11) is reformulated as

$$R_{3i} = \begin{cases} \max\{0, E_i \exp(\alpha_3 - \beta_3 E_i - \zeta_3/(w_i - \xi_3) + \epsilon_i\} & \text{if } w_i > \xi_3 \\ \max\{0, \epsilon_i\} & \text{if } w_i \leq \xi_3 \end{cases}$$

$$i = 67...83 \qquad (6.12)$$

The parameters of model (6.12) are estimated by maximum likelihood. If all the values of an R_3 data set are non-zero (as for Black Brows Beck), the maximum likelihood estimates are identical to the nonlinear least squares estimates.

The reformulation makes inference more complicated in two ways. First, model (6.12) can have censored observations, since when $R_{3i} = 0$, the error ϵ_i is effectively censored. Secondly, the model has a change point at $w = \xi_3$. This considerably increases the nonlinear behaviour of the model (see Chapter 7). For example, a simulated distribution of $\hat{\xi}_3$ is very skewed and has at least two modes, one in the region of the "true value" and the change point $\hat{\xi}_3 = w_{83} = 1.64$ and one in the region of the change point $\hat{\xi}_3 = w_{76} = 1.85$ (Fig. 6.13). (Model (6.12) and some alternative models are discussed in more detail in Section 6.3.3).

It was not possible to find a parameterisation of (6.12) which had approximately close-to-linear behaviour. The parameter transformation that was useful for the R_2 data set,

$$\gamma_{\omega 1} = \exp(-\zeta/(\omega_1 - \xi))$$

with $\omega_1 = 1.64$, can not be used in this case, because it is not well defined in the 95% region of interest (ie it is not defined at $\xi_3 = 1.64$). Even if the definition is modified so that $\gamma_{\omega 1} = 0$ when $\xi_3 = 1.64$, the transformation is still discontinuous at $\xi_3 = 1.64$. If $\omega_1 > 2.1$ say, the parameter $\gamma_{\omega 1}$ is well defined and continuous in the region of interest; however, the transformation is of little use since the distribution of $\hat{\gamma}_{\omega 1}$, although approximately unimodal, is extremely skewed. Thus, although model (6.12) is a good description of the underlying relationship between

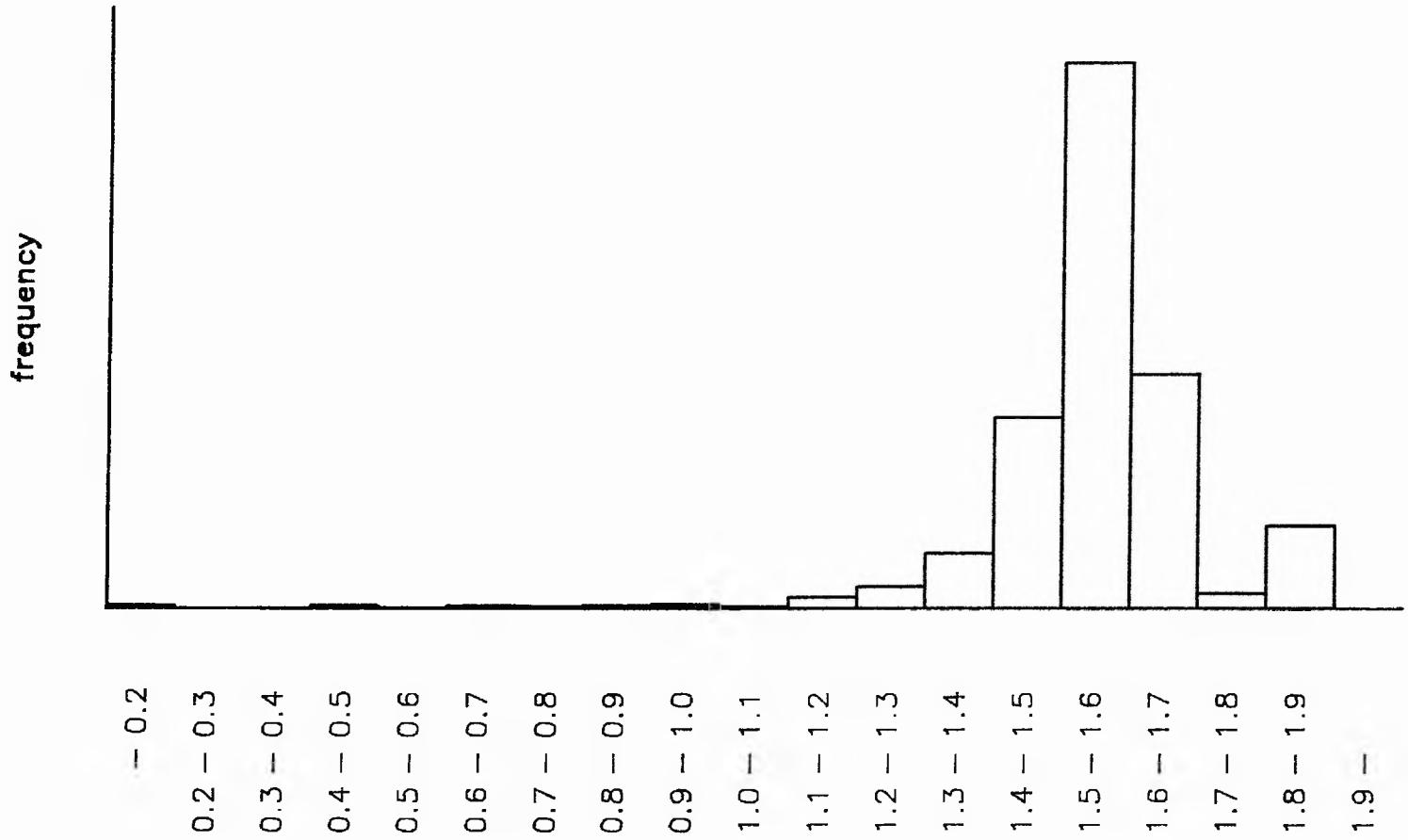


Fig. 6.13 Simulated distribution of $\hat{\xi}_3$.

R_3, E and w, the nonlinear behaviour of the model is so large that it is not possible to express parameter confidence regions in a simple form. Instead, approximate 95% confidence intervals for α_3, β_3, ζ_3, ξ_3 based on a simulation are given in Table 6.7.

6.3.2 Survival Between t_2 and t_3

To investigate the survival rate between t_2 and t_3 models (6.7) and (6.12) are combined to give

$$R_{ji} = \begin{cases} \max\{0, E_i \exp(\alpha_j - \beta_j E_i - \zeta_j/(w_i - \xi_j)) + \epsilon_{ji}\} & \text{if } w_i > \xi_j \\ \max\{0, \epsilon_{ji}\} & \text{if } w_i \le \xi_j \end{cases} \qquad j = 2,3, \; i = 67 \ldots 83 \qquad (6.13)$$

where ϵ_{ji} and ϵ_{kl} are independent if $i \neq l$ and where

$$\begin{pmatrix} \epsilon_{2i} \\ \epsilon_{3i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_2^2 & \rho\sigma_2\sigma_3 \\ \rho\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix} \right), \quad i = 67 \ldots 83$$

The likelihood of (6.13) is maximised with

i)	all the parameters unconstrained,	$-2\log L = 134.86$,
ii)	the constraints $\zeta_2 = \zeta_3$, $\xi_2 = \xi_3$,	$-2\log L = 135.85$,
iii)	as ii) and the constraint $\beta_2 = \beta_3$,	$-2\log L = 136.03$,
iv)	as iii) and the constraint $\rho = 0$,	$-2\log L = 138.35$.

Hence $\zeta_2 = \zeta_3$ and $\xi_2 = \xi_3$ ($-2\log\lambda = 0.99$, $\chi^2(2;0.05) = 5.99$, $p > 0.05$), so the "drought effect" at t_3 is the same as the "drought effect" at t_2. Thus, there is no residual mortality between t_2 and t_3 due to the summer droughts. Also, $\beta_2 = \beta_3$

($-2\log\lambda = 0.18, \chi^2(1;0.05) = 3.84$, $p > 0.05$), so survival between t_2 and t_3 is not density dependent. Finally, ρ is not significantly different from zero ($-2\log\lambda = 2.32, \chi^2(1;0.05) = 3.84$, $p > 0.05$), so the errors $\{\varepsilon_{2i}\}$ and $\{\varepsilon_{3i}\}$ are approximately uncorrelated. Thus R_2 and R_3 are well described by the model

$$R_{ji} = \begin{cases} \max\{0, E_i \exp(\alpha_j - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{ji}\} & \text{if } w_i > \xi_2 \\ \max\{0, \epsilon_{ji}\} & \text{if } w_i \le \xi_2 \end{cases}$$

$$j = 2,3, \quad i = 67...83 \qquad (6.14)$$

where $\{\varepsilon_{2i}\}$ are NID$(0,\sigma_2^2)$, $\{\varepsilon_{3i}\}$ are NID$(0,\sigma_3^2)$ and where $\{\varepsilon_{2i}\}$ are independent of $\{\varepsilon_{3i}\}$. Parameter estimates are given in Table 6.9.

The change points at $\xi_3 = 1.64$ and $\xi_3 = 1.85$ are in the region of interest when all the parameters are unconstrained. This is likely to cause large intrinsic nonlinearity so the result of the first likelihood ratio test is checked by a simulation. 100 R_2 and R_3 data sets are generated using the estimates from ii) above. The maximum likelihoods of the observations under constraints i) and ii) are then compared. The simulated critical value for rejecting H_0: $\zeta_2 = \zeta_3, \xi_2 = \xi_3$ in favour of H_1: $\zeta_2 \neq \zeta_3, \xi_2 \neq \xi_3$ at the 95% significance level is $\chi^2_{sim}(0.05) = 6.9$. Since $-2\log\lambda = 0.99$, the simulation confirms that $\zeta_2 = \zeta_3$ and $\xi_2 = \xi_3$. Once it is assumed that $\zeta_2 = \zeta_3$ and $\xi_2 = \xi_3$, the change points at $\xi_3 = 1.64$ and $\xi_3 = 1.85$ are no longer in the region of interest. Hence, the intrinsic nonlinearity of the model is considerably reduced, so simulations are not used to confirm the other likelihood ratio tests.

TABLE 6.9

Parameter estimates for model (6.14)

Parameter	Estimate	s.e.	Correlations				
α_2	-2.353	0.031	1.00				
α_3	-3.249	0.053	0.53	1.00			
β_2	$2.744*10^{-4}$	$0.051*10^{-4}$	0.76	0.44	1.00		
$\gamma_{\omega 1}$	0.3578	0.0339	-0.05	-0.06	-0.05	1.00	
$\gamma_{\omega 2}$	0.8523	0.0229	-0.60	-0.38	-0.09	-0.04	1.00

Parameter	Estimate	95% confidence limits	
ζ_2	0.120	0.097	0.176
ξ_2	1.529	1.465	1.572

Variance	Estimate
σ_2^2	12.58
σ_3^2	36.82

A simulation study of model (6.14) shows that $\hat{\alpha}_2$, $\hat{\alpha}_3$ and $\hat{\beta}_2$ are approximately normally distributed, but that the distributions of $\hat{\zeta}_2$ and $\hat{\xi}_2$ are very skewed (Table 6.10). There are no change points in the region of interest, so the parameter transformation

$$\gamma_{\omega j} = \exp(-\zeta_2/(\omega_j - \xi_2)), \quad j = 1,2,$$

with $\omega_1 = 1.64$ and $\omega_2 = 2.28$ can be used. The distribution of $\hat{\gamma}_{\omega 1}$ is approximately normal and the distribution of $\hat{\gamma}_{\omega 2}$ is much closer-to-normal than those of $\hat{\zeta}_2$ and $\hat{\xi}_2$ (Table 6.10), so it is concluded that (α_2, α_3, β_2, $\gamma_{\omega 1}$, $\gamma_{\omega 2}$) is a much better parameterisation of (6.14) than (α_2, α_3, β_2, ζ_2, ξ_2). Approximate standard errors and correlations (obtained from the simulation) are given in Table 6.9.

The relationship between R_2 and R_3 is discussed in Section 6.6.

6.3.3 Other Models

Other approaches to modelling the R_2 and R_3 data can also be considered. In particular, a multiplicative error can be used instead of an additive error. The models

$$\begin{aligned} R_{2i} &= E_i \exp(\alpha_2 - \beta_2 E_i - \zeta_2/(w_i - \xi_2) + \epsilon_{2i}) && \text{if } w_i > \xi_2 \\ &\quad 0 && \text{if } w_i \le \xi_2 \\ & && i = 67 \ldots 83, \end{aligned} \qquad (6.15)$$

TABLE 6.10

Tests of normality of the m.l.e.s of model (6.14)

Parameter	α_2	α_3	β_2
T_3 - skewness stat	1.28	-1.30	-0.69
T_4 - excess kurtosis stat	-0.04	0.35	-0.10

Parameter	ζ_2	ξ_2	$\gamma_{\omega 1}$	$\gamma_{\omega 2}$
T_3 - skewness stat	8.01**	-8.30**	0.60	-3.41**
T_4 - excess kurtosis stat	7.50**	3.84**	-0.60	3.23**

** indicates serious nonlinear behaviour

where $\{\epsilon_{2i}\}$ are NID$(0,\sigma_2^2)$, and

$$R_{3i} = \begin{cases} E_i \exp(\alpha_3 - \beta_3 E_i - \zeta_3/(w_i - \xi_3) + \epsilon_{3i}) & \text{if } w_i > \xi_3 \\ 0 & \text{if } w_i \le \xi_3 \end{cases} \qquad i = 67 \ldots 83, \qquad (6.16)$$

where $\{\epsilon_{3i}\}$ are NID$(0,\sigma_3^2)$, are excellent fits to the R_2 and R_3 data sets. An examination of the residuals gives no evidence that the multiplicative error structure is not appropriate. Thus, there is not sufficient data to determine whether an additive or a multiplicative error is more appropriate (or if either are appropriate).

Inference depends on the error structure used. For example, if the multiplicative model (6.16) is assumed to be the "true" model, the "true" value of ξ_3 must be less than 1.64 = min$\{w_i\}$, since all the observed R_3 values are strictly positive. However, if the additive model (6.12) is the true model, ξ_3 could be greater than 1.64; indeed, in a simulation study, 20% of the estimates of ξ_3 were greater than 1.64.

However, in terms of the dynamics of the trout population, it is more important to compare parameters between different life stages (ie to determine whether survival between the life stages is density dependent) than to make inferences about the parameters within a life stage. Hence, models (6.15) and (6.16) are combined to give

$$R_{ji} = \begin{cases} E_i \exp(\alpha_j - \beta_j E_i - \zeta_j/(w_i - \xi_j) + \epsilon_{ji}) & \text{if } w_i > \xi_j \\ 0 & \text{if } w_i \le \xi_j \end{cases} \qquad i = 67 \ldots 83, \quad j = 2,3$$

where ϵ_{ji} and ϵ_{kl} are independent if $i \neq l$ and where

$$\begin{pmatrix} \epsilon_{2i} \\ \epsilon_{3i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_2^2 & \rho\sigma_2\sigma_3 \\ \rho\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix} \right), \quad i = 67...83$$

Maximising the likelihood with

i) all the parameters unconstrained, $-2\log L = -137.08$,

ii) the constraints $\zeta_2 = \zeta_3$, $\xi_2 = \xi_3$, $\beta_2 = \beta_3$, $-2\log L = -130.98$,

gives $\zeta_2 = \zeta_3$, $\xi_2 = \xi_3$ and $\beta_2 = \beta_3$ ($-2\log\lambda = 6.10, \chi^2(3;0.05) = 7.81$, $p > 0.05$). Hence, the choice of error structure does not affect the inference at the 95% significance level that survival between t_2 and t_3 is density independent and that there is no residual mortality between t_2 and t_3 due to the summer droughts.

The multiplicative error model (M) appears to overcome the problems of inference that arise when the additive error model (A) is applied to the R_3 data set, because under (M)

i) the number of recruits must be non-negative,

ii) the "true" value of ξ_3 must be less than 1.64, (since the R_3 observations are strictly positive,) so the change point at $w = \xi_3$ can be ignored.

However, (A) has one major advantage over (M) which relates to the form of the two models when the expected number of recruits is low. Figure 6.14 shows the expected number of recruits for both models as a function of w for fixed E; also shown are approximate 95% confidence limits for an observation R. Under (A), the variance of R (given E and w) is similar for all values of w (although it decreases

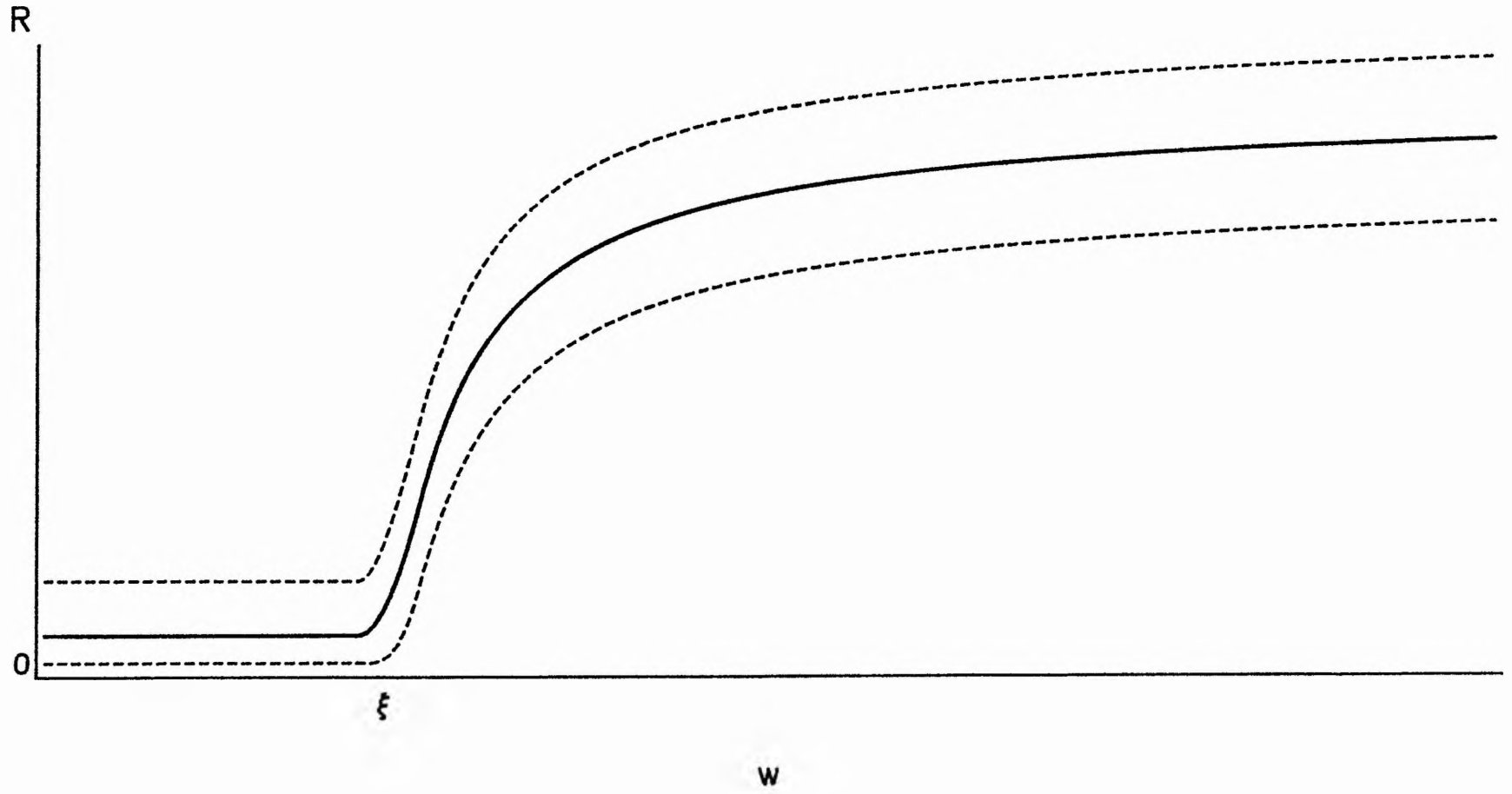


Fig. 6.14a The number of recruits as a function of w — additive error model.

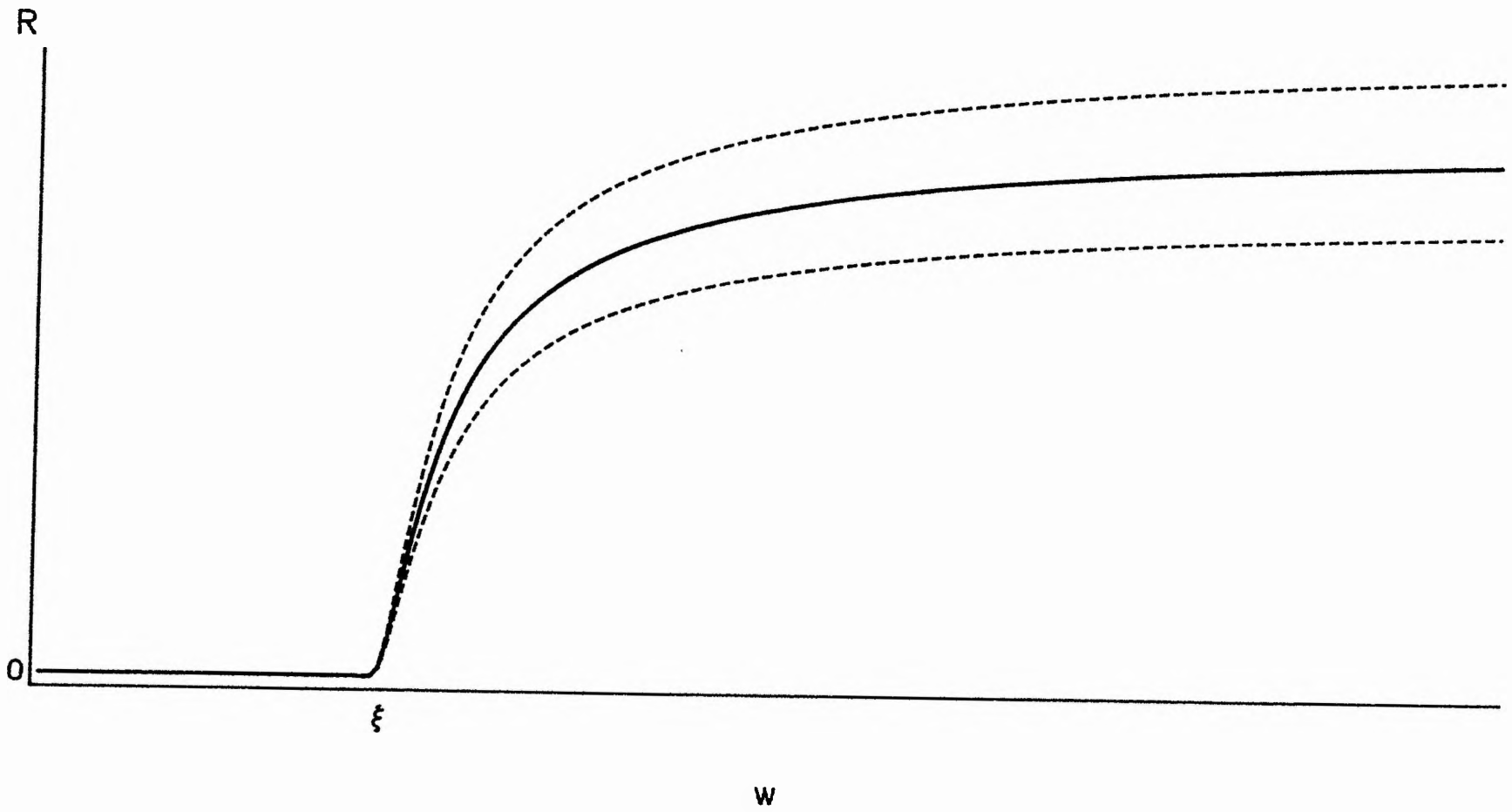


Fig. 6.14b The number of recruits as a function of w — multiplicative error model.

slightly as w decreases towards ξ). Under (M), the variance of R decreases as w decreases and is zero if $w \leq \xi$. Thus, (A) is the more conservative of the two models, in the sense that recruitment in the drought years is more variable under (A) than under (M). This is particularly desirable, because the form of the drought term

$$\exp(-\zeta/(w - \xi)) \tag{6.17}$$

is based on the observed survival rate in only a small number of years. With such a small amount of data, it is inevitable that (6.17), or any other way of incorporating a drought effect, will be a very simplistic representation of the "true" response of the population to drought (although (6.17) appears to work very well). Also, there is little information available on the variability of survival in a drought year and none available on survival when $w < 1.64$. This "uncertainty" about the "true" response to drought (due to the lack of data) is accommodated better by (A) than (M) through the greater variability of an observation in a drought year under (A). Hence, the additive error model is used in preference to the multiplicative error model.

A second advantage of (A) arises because the number of recruits R is integer valued. Under (A), this can effectively be ignored since the variance of R is "large" for all values of w; ie a large number of integer values fall within the 95% confidence limits in Figure 6.14a. However, under (M), the variance of R becomes very small when w is close to ξ and the discrete properties of R must be taken into account. The multiplicative model could be reformulated as, say

$$R_i = \begin{cases} [E_i \exp(\alpha - \beta E_i - \zeta/(w_i - \xi) + \epsilon_i) + 0.5] & \text{if } w_i > \xi \\ 0 & \text{if } w_i \leq \xi \end{cases}$$
$$i = 67 \ldots 83.$$

where $[x]$ denotes the integer part of x. However, estimation and inference then becomes much more complicated.

The non-negative, integer valued number of recruits R could be more sensibly accommodated by assuming that R is Poisson distributed with mean

$$\begin{cases} E_i \exp(\alpha - \beta E_i - \zeta/(w_i - \xi)) & \text{if } w_i > \xi \\ 0 & \text{if } w_i \leq \xi \end{cases}$$
$$i = 67 \ldots 83.$$

However, like the multiplicative model, the Poisson model gives very high weight to the observations in the drought years. Therefore, it is not considered further.

6.4 R_4 - 1+ PARR (AUGUST/SEPTEMBER)

6.4.1 Survival Between t_0 and t_4

The model

$$R_{4i} = \begin{cases} \max\{0, E_i \exp(\alpha_4 - \beta_4 E_i - \zeta_4/(w_i - \xi_4)) + \epsilon_i\} & \text{if } w_i > \xi_4 \\ \max\{0, \epsilon_i\} & \text{if } w_i \leq \xi_4 \end{cases}$$
$$i = 67 \ldots 83$$

where $\{\epsilon_i\}$ are NID$(0, \sigma^2)$, is an unsatisfactory description of the R_4 data set, because it does not account for mortality due to drought in the second summer of the life

cycle. For example, survival of the 1967 year class between t_3 and t_4 is affected by the drought of 1968, etc. Plotting the residual corresponding to the ith year class against w_{i+1} shows that the residuals become larger and more negative as w_{i+1} decreases (ie as the severity of a drought increases) (Fig. 6.15). Thus, $\{w_{i+1}, i = 67...83\}$ are used as measures of drought in the second summer of the life cycle.

The model

$$R_{4i} = \begin{cases} \max\{0, E_i \exp(\alpha_4 - \beta_4 E_i - \zeta_4/(w_i - \xi_4) - \zeta_4'/(w_{i+1} - \xi_4')) + \epsilon_i\} & \text{if } w_i > \xi_4 \text{ and } w_{i+1} > \xi_4' \\ \max\{0, \epsilon_i\} & \text{otherwise} \end{cases}$$

$$i = 67...83 \qquad (6.18)$$

where $\{\epsilon_i\}$ are NID$(0, \sigma^2)$, is an excellent description of the R_4 data. The drought effect in the second summer of the life cycle is well accounted for by the inclusion of the second drought term ($-2\log\lambda = 17.42, \chi^2(1;0.05) = 3.84, \chi^2_{sim}(0.05) = 7.22$, $p < 0.05$). The residuals give no evidence that the form of the error term is not appropriate and reveal no systematic departures when plotted against either $\{w_i\}$ or $\{w_{i+1}\}$. Parameter estimates are given in Table 6.11. Model (6.18) has very large nonlinear behaviour. For example, the maximum intrinsic curvature (which neglects the additional nonlinearity due to the change points) is $\Gamma^N = 1.48$, compared to a "critical value" of $1/(2\sqrt{F}) = 0.28$. Hence, only approximate confidence intervals obtained from a simulation are given for each parameter.

6.4.2 Survival Between t_2 and t_4

The survival rate between t_2 and t_4 is investigated by combining (6.14) and (6.18) to give a model relating R_2, R_3 and R_4 to E and w;

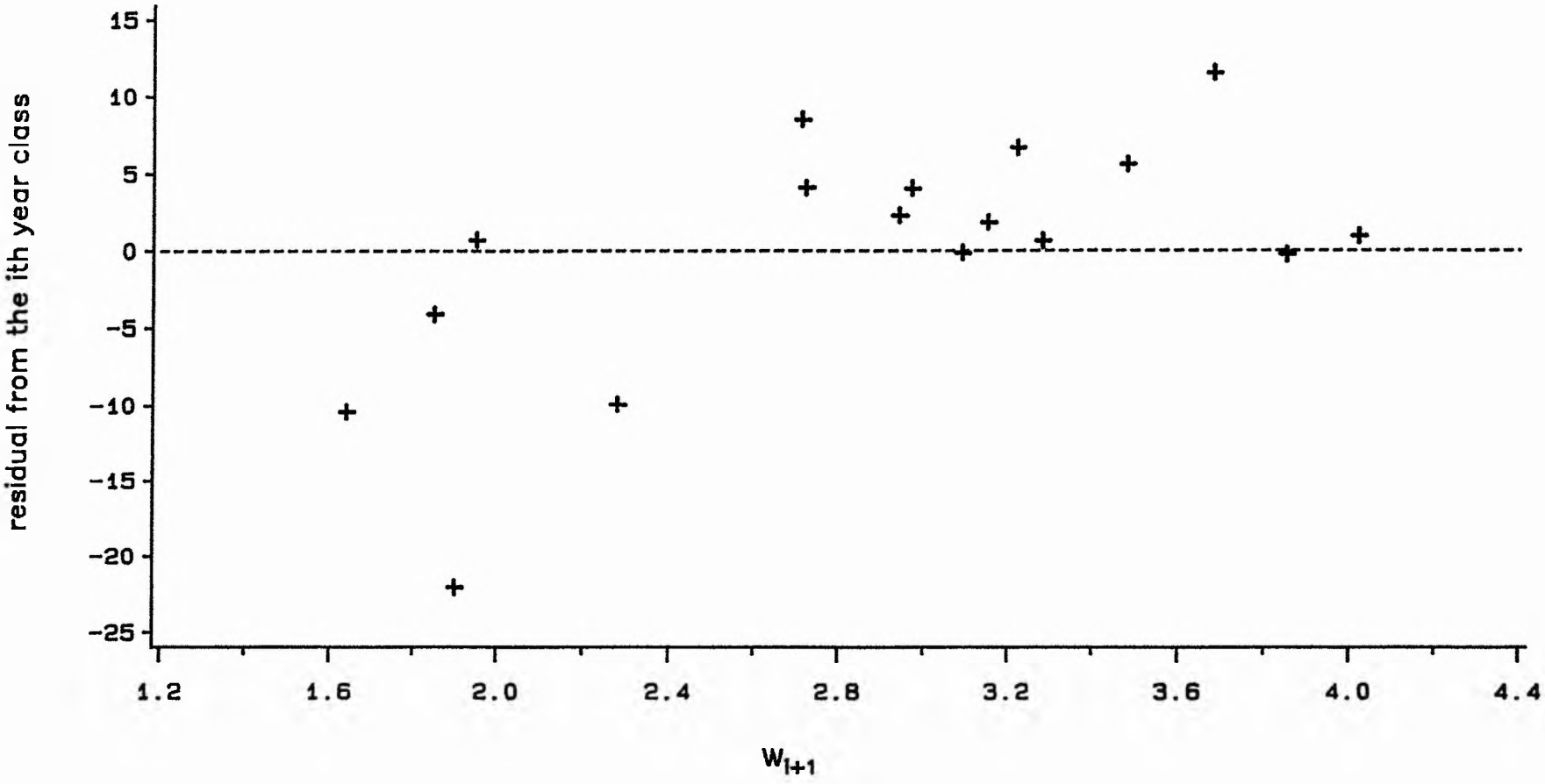


Fig. 6.15 Residuals against $\{w_{i+1}\}$.

TABLE 6.11

Parameter estimates for model (6.18)

Parameter	Estimate	95% confidence limits	
α_4	-3.055	-3.460	-2.220
β_4	$2.864*10^{-4}$	$2.380*10^{-4}$	$3.320*10^{-4}$
ζ_4	0.045	$1.1*10^{-5}$	0.235
ξ_4	1.627	1.380	1.860
ζ_4'	0.568	0.085	7.913
ξ_4'	1.198	-0.750	1.630

rss = 370.46 on 11 d.f.
s^2 = 33.68

$$R_{ji} = \begin{cases} \max\{0, E_i \exp(\alpha_j - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{ji}\} & \text{if } w_i > \xi_2 \\ \max\{0, \epsilon_{ji}\} & \text{if } w_i \leq \xi_2 \end{cases}$$
$$j = 2,3.$$

$$R_{4i} = \begin{cases} \max\{0, E_i \exp(\alpha_4 - \beta_4 E_i - \zeta_4/(w_i - \xi_4) - \zeta_4'/(w_{i+1} - \xi_4')) + \epsilon_{4i}\} & \text{if } w_i > \xi_4 \text{ and } w_{i+1} > \xi_4' \\ \max\{0, \epsilon_{4i}\} & \text{otherwise} \end{cases}$$
$$i = 67...83, \quad (6.19)$$

where ϵ_{2i}, ϵ_{3i}, ϵ_{4i} are independent of ϵ_{2k}, ϵ_{3k}, ϵ_{4k} if $i \neq k$ and with

$$\begin{pmatrix} \epsilon_{2i} \\ \epsilon_{3i} \\ \epsilon_{4i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_2^2 & 0 & 0 \\ 0 & \sigma_3^2 & \rho\sigma_3\sigma_4 \\ 0 & \rho\sigma_3\sigma_4 & \sigma_4^2 \end{pmatrix} \right), \quad i = 67...83$$

The likelihood is maximised with

i) all the parameters unconstrained, $-2\log L = 195.55$,

ii) the constraints $\zeta_4 = \zeta_2$, $\xi_4 = \xi_2$, $-2\log L = 199.05$,

iii) as ii) and the constraint $\beta_4 = \beta_2$, $-2\log L = 199.19$,

iv) as iii) and the constraint $\rho = 0$, $-2\log L = 211.45$.

Hence $\zeta_4 = \zeta_2$, $\xi_4 = \xi_2$ ($-2\log\lambda = 3.50, \chi^2(2;0.05) = 5.99, \chi^2_{sim}(0.05) = 9.72$, $p > 0.05$) so, as would be expected, a drought in the first summer of the life cycle does not affect survival in the second summer of the life cycle. Also, $\beta_4 = \beta_2$ ($-2\log\lambda = 0.14, \chi^2(1;0.05) = 3.84, \chi^2_{sim}(0.05) = 5.75$, $p > 0.05$), so survival between t_2 and t_4 is not density dependent. Finally, $\rho \neq 0$ ($-2\log\lambda = 12.26, \chi^2(1;0.05) = 3.84, \chi^2_{sim}(0.05) = 3.87$, $p < 0.05$), so the errors $\{\epsilon_{3i}\}$ and $\{\epsilon_{4i}\}$ are not independent. The correlation between the errors is positive, so if, for example, the survival rate

of a year class between t_0 and t_3 is greater than expected, the survival rate between t_0 and t_4 will also tend to be greater then expected. Combining the results above enables model (6.19) to be rewritten as

$$
\begin{aligned}
R_{ji} = \quad & \max\{0, E_i \exp(\alpha_j - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{ji}\} && \text{if } w_i > \xi_2 \\
& \max\{0, \epsilon_{ji}\} && \text{if } w_i \leq \xi_2 \\
& && j = 2,3, \\
R_{4i} = \quad & \max\{0, E_i \exp(\alpha_4 - \beta_2 E_i - \zeta_2/(w_i - \xi_2) - \zeta_4'/(w_{i-1} - \xi_4')) + \epsilon_{4i}\} \\
& && \text{if } w_i > \xi_2 \text{ and } w_{i-1} > \xi_4' \\
& \max\{0, \epsilon_{4i}\} && \text{otherwise} \\
& && i = 67 \ldots 83, \qquad (6.20)
\end{aligned}
$$

Parameter estimates are given in Table 6.12.

6.5 EGG PRODUCTION

The egg production of a year class is estimated from the number of spawning females in the year class and the time at which they spawn (cf Section 5.2.1). The spawning season is divided into three periods: before 14 November, 14–25 November and after 25 November. Let μ_j be the mean number of eggs in a redd cut in period j, j = 1,2,3, estimated to be μ_1 = 1550, μ_2 = 1034, μ_3 = 722 (Section 5.2.1). Let z_{3ij} and z_{4ij} be the number of 3+ and 4+ females respectively in the ith year class that spawn in period j; for example, $z_{3,67j}$ is the number of 3+ females that spawn before 14th November in 1970 etc. Then the egg production of the ith year class is estimated to be

TABLE 6.12

Parameter estimates for model (6.20)

Parameter	Estimate	95% confidence limits	
α_2	-2.353	-2.406	-2.297
α_3	-3.250	-3.342	-3.159
α_4	-3.116	-3.280	-2.820
β_2	$2.750*10^{-4}$	$2.661*10^{-4}$	$2.841*10^{-4}$
ζ_2	0.116	0.083	0.166
ξ_2	1.532	1.474	1.570
ζ_4'	0.361	0.125	1.164
ξ_4'	1.451	0.813	1.697

Variance	Estimate
σ_2^2	12.59
σ_3^2	36.94
σ_4^2	29.10
ρ	0.741

$$E_i^* = \sum_{j=1}^{3} (z_{3ij} + z_{4ij})\mu_j, \quad i = 67...83$$

Ideally, the egg production should be investigated through the distribution of $\{z_i = (z_{3i1}, z_{3i2}, z_{3i3}, z_{4i1}, z_{4i2}, z_{4i3}), i = 67...83\}$. However, there is not enough data to attempt an investigation of this sort. Instead, nonlinear regression models are used to model E_i^* as a function of E_i, w_i and w_{i+1}, $i = 67...83$.

The model

$$E_i^* = \begin{cases} \max\{0, E_i \exp(\alpha_5 - \beta_5 E_i - \zeta_5/(w_i - \xi_5) - \zeta_5'/(w_{i+1} - \xi_5')) + \epsilon_i\} & \text{if } w_i > \xi_5 \text{ and } w_{i+1} > \xi_5' \\ \max\{0, \epsilon_i\} & \text{otherwise} \end{cases}$$
$$i = 67...83 \qquad (6.21)$$

where $\{\epsilon_i\}$ are NID$(0,\sigma^2)$ is fitted by maximum likelihood with

i) the constraints $\zeta_5 = \zeta_5' = 0$, $-2\log L = 273.43$,

ii) the constraint $\zeta_5' = 0$, ζ_5 unconstrained, $-2\log L = 265.13$,

iii) ζ_5, ζ_5' unconstrained. $-2\log L = 253.98$.

The inclusion of each drought term significantly improves the fit of the model (H_0: $\zeta_5 = \zeta_5' = 0$ vs H_1: $\zeta_5 \neq 0$, $\zeta_5' = 0$, $-2\log\lambda = 8.30$, $\chi^2(1;0.05) = 3.84$, $\chi^2_{sim}(0.05) = 5.66$, $p < 0.05$; H_0: $\zeta_5 \neq 0$, $\zeta_5' = 0$ vs H_1: $\zeta_5 \neq 0$, $\zeta_5' \neq 0$, $-2\log\lambda = 11.15$, $\chi^2(1;0.05) = 3.84$, $\chi^2_{sim}(0.05) = 5.68$, $p < 0.05$). An examination of the residuals reveals no lack of fit of the full model. Further, the variance σ^2 in the full model is estimated to be $(1323)^2$. Since one extra spawning female can increase egg production by between 700 and 1500 eggs, this effectively means that the number

of spawning females can be "predicted" with an error of approximate standard deviation between one and two. This is as accurate a level of prediction as one might expect to achieve and suggests that, at this comparatively simple level of modelling, there is "no more variability to be explained". (The level of accuracy is also a plausible one since, on average, there were approximately four spawning females in each year class). Hence, model (6.21) is a good description of the egg production data. Parameter estimates are given in Table 6.13.

The 95% parameter confidence limits given in Table 6.13 are very large, particularly in comparison with the confidence limits of the parameters β_2, ζ_2, ξ_2 in model (6.20) (Table 6.12). Hence, a simple method of comparing the measures of density dependence and the "drought effects" at t_4 and the time of spawning is to assume that β_2, ζ_2, ξ_2 are known a priori and are given by their estimates in Table 6.12 and that the relationship between R_4, E^* and E can be described by the model

$$R_{4i} = \begin{cases} \max\{0, E_i \exp(\alpha_4 - 2.750*10^{-4}E_i - 0.1160/(w_i - 1.532) - \zeta_4'/(w_{i-1} - \xi_4')) + \epsilon_{4i}\} & \text{if } w_i > \xi_4 \text{ and } w_{i-1} > \xi_4' \\ \max\{0, \epsilon_{4i}\} & \text{otherwise} \end{cases}$$

$$E_i^* = \begin{cases} \max\{0, E_i \exp(\alpha_5 - \beta_5 E_i - \zeta_5/(w_i - \xi_5) - \zeta_5'/(w_{i-1} - \xi_5')) + \epsilon_{5i}\} & \text{if } w_i > \xi_5 \text{ and } w_{i-1} > \xi_5' \\ \max\{0, \epsilon_{5i}\} & \text{otherwise} \end{cases}$$

$$i = 67 \ldots 83, \qquad (6.22)$$

with $\epsilon_{ji}, \epsilon_{kl}$ independent if $i \neq l$ and with

TABLE 6.13

Parameter estimates for model (6.21)

Parameter	Estimate	95% confidence limits	
α_5	1.925	1.544	3.773
β_5	$3.272*10^{-4}$	$2.492*10^{-4}$	$4.039*10^{-4}$
ζ_5	0.183	$3*10^{-6}$	22.74
ξ_5	1.523	-3.907	1.930
ζ_5'	0.097	$1*10^{-5}$	8.275
ξ_5'	1.787	-1.005	2.091

rss = $1.927*10^7$ on 11 d.f.
s^2 = $1.752*10^6$

$$\begin{pmatrix}\epsilon_{4i}\\ \epsilon_{5i}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_4^4 & \rho\sigma_4\sigma_5\\ \rho\sigma_4\sigma_5 & \sigma_5^4\end{pmatrix}\right), \quad i = 67...83$$

The likelihood of (6.22) is maximised with

i) all the parameters unconstrained, $-2\log L = 89.71$,

ii) the constraints $\beta_5 = \beta_2 = 2.750*10^{-4}$

$\zeta_5 = \zeta_2 = 0.1160$

$\xi_5 = \xi_2 = 1.532$

and $\zeta_5' = \zeta_4'$, $\xi_5' = \xi_4'$, $-2\log L = 97.01$.

Hence $(\beta_5, \zeta_5, \xi_5, \zeta_5', \xi_5') = (\beta_2, \zeta_2, \xi_2, \zeta_4', \xi_4')$ ($-2\log\lambda = 7.30$, $\chi^2(5;0.05) = 11.07$, $\chi^2_{sim}(0.05) = 14.22$, $p > 0.05$). In particular, $\beta_5 = \beta_2$, so "survival" between t_2 and the time of spawning is density independent.

6.6 DISCUSSION

Survival is density dependent between t_0 and t_1 and between t_1 and t_2 (Section 6.2.7). The relationships between R_1 and E and between R_2 and E are well described by models (6.6) and (6.7) and are discussed in Chapter 4 and Section 6.2 respectively. The relationship between R_1 and R_2 is much more complicated (cf equation (6.8)) and is investigated in more detail in Chapter 9.

After t_2, survival is density independent (Sections 6.3.2, 6.4.2, 6.5) and it is usually possible to describe the relationship between one life stage and the next by a comparatively simple model. For example, consider the life stages R_2 and R_3. The relationship between R_2, R_3 and E is well described by (6.14). For simplicity, it is

assumed that $w_i > \xi_2$, $i = 67...83$, is known a priori. Also, given the parameter values in Table 6.9, the probability of obtaining zero trout in any year class at time t_2 is negligible. Hence (6.14) can be written

$$R_{2i} = E_i \exp(\alpha_2 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{2i}$$

$$R_{3i} = \max\{0, E_i \exp(\alpha_3 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{3i}\}, \quad i = 67...83 \qquad (6.23)$$

where $\{\epsilon_{2i}\}$ are NID$(0, \sigma_2^2)$, $\{\epsilon_{3i}\}$ are NID$(0, \sigma_3^2)$ and $\{\epsilon_{2i}\}$ are independent of $\{\epsilon_{3i}\}$. The measure of density dependent survival is the same for both life stages, so E can be eliminated from (6.23) to give

$$R_{3i} = \max\{0, R_{2i} \exp(\alpha_3 - \alpha_2) + \epsilon_{23i}\}, \quad i = 67...83$$

where

$$\epsilon_{23i} = \epsilon_{3i} - \epsilon_{2i} \exp(\alpha_3 - \alpha_2), \quad i = 67...83.$$

The errors $\{\epsilon_{23i}\}$ are NID$(0, \sigma_3^2 + \sigma_2^2 \exp(2(\alpha_3 - \alpha_2)))$. Thus, with a suitable reparameterisation, the relationship between R_2 and R_3 can be described by the model

$$R_{3i} = \max\{0, KR_{2i} + \epsilon_i\}, \quad i = 67..83 \qquad (6.24)$$

where (ε_i) are NID$(0,\sigma^2)$. The arguments leading to (6.24) are corroborated by fitting (6.24) directly to the R_2 and R_3 data sets; model (6.24) is a good fit to the data (Fig. 6.16) and an examination of the residuals reveals no evidence of lack of fit or that the form of the error term is inappropriate.

The relationship between R_3 and R_4 is more complicated because there is a non-negligible probability of obtaining zero trout at t_3. From (6.20), the relationship between R_3, R_4 and E is given by

$$R_{3i} = \max\{0, E_i \exp(\alpha_3 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_{3i}\}$$

$$R_{4i} = \max\{0, E_i \exp(\alpha_4 - \beta_2 E_i - \zeta_2/(w_i - \xi_2) - \zeta_4'/(w_{i+1} - \xi_4')) + \epsilon_{4i}\}$$

$$i = 67 \ldots 83 \qquad (6.25)$$

where ε_{ji} and ε_{kl} are independent if $i \neq l$ and where

$$\begin{pmatrix} \epsilon_{3i} \\ \epsilon_{4i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 & \rho\sigma_3\sigma_4 \\ \rho\sigma_3\sigma_4 & \sigma_4^2 \end{pmatrix}\right), \quad i = 67 \ldots 83$$

If $R_{3i} \neq 0$, E can be eliminated from (6.25) to give

$$R_{4i} = \max(0, R_{3i} \exp(\alpha_4 - \alpha_3 - \zeta_4'/(w_{i+1} - \xi_4')) + \epsilon_{34i}\}$$

where

$$\epsilon_{34i} = \epsilon_{4i} - \epsilon_{3i} \exp(\alpha_4 - \alpha_3 - \zeta_4'/(w_{i+1} - \xi_4'));$$

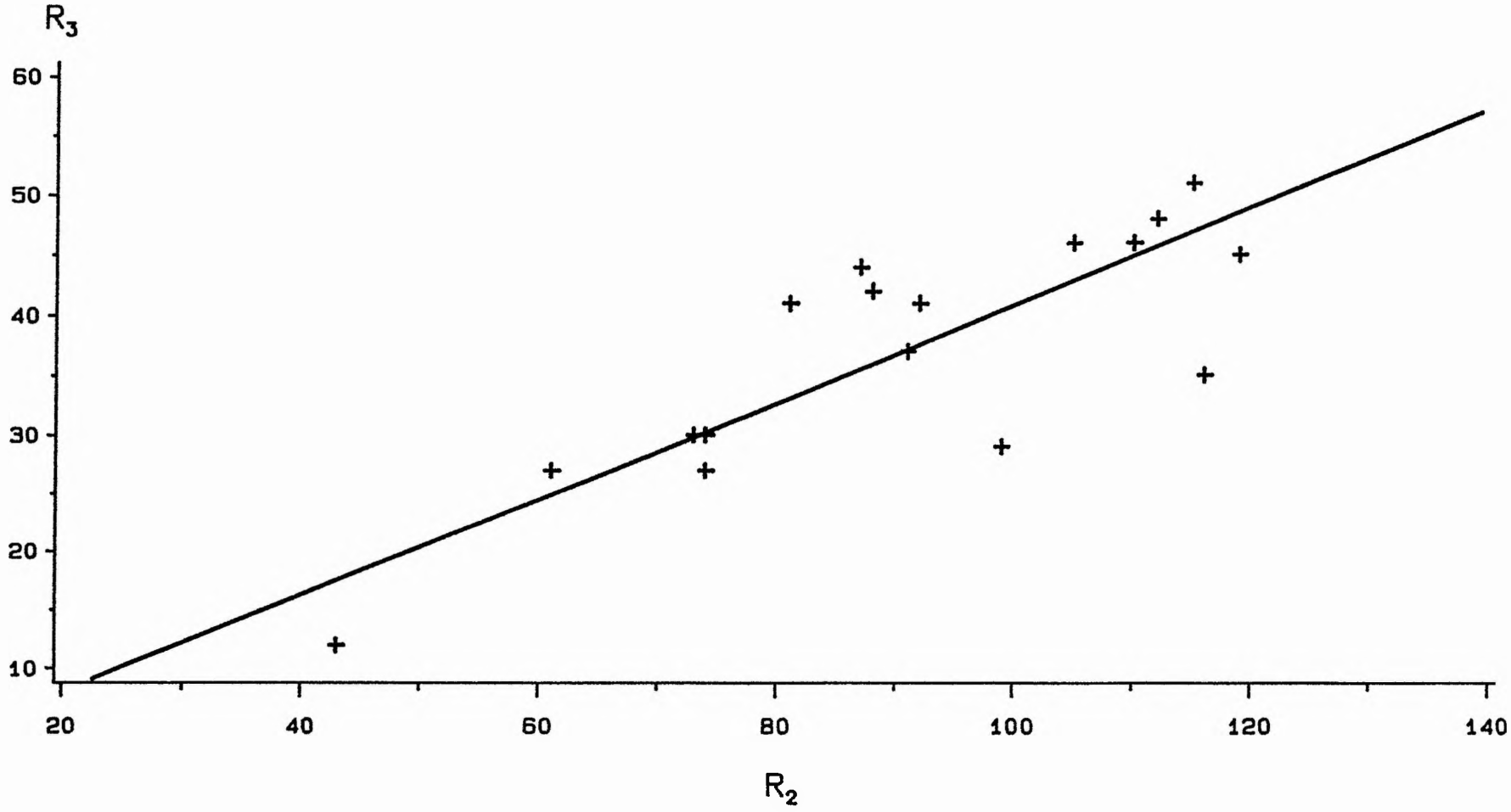


Fig. 6.16 Model (6.24) fitted to the R_2 and R_3 data sets.

the error ϵ_{34i} is normally distributed with zero mean and variance

$$\sigma_4^2 - 2\rho\sigma_3\sigma_4\exp(\alpha_4 - \alpha_3 - \zeta_4'/(w_{i+1} - \xi_4')) + \sigma_3^2\exp(2(\alpha_4 - \alpha_3 - \zeta_4'/(w_{i+1} - \xi_4'))$$

However, if $R_{3i} = 0$, E_i can not be eliminated from (6.25) and

$$R_{4i} = \max\{0, E_i\exp(\alpha_4 - \beta_2 E_i - \zeta_2/(w_i - \xi_2) - \zeta_4'/(w_{i+1} - \xi_4') + \epsilon_{4i}\} \qquad (6.26)$$

There are two points to note about this relationship. First, it is not inconsistent that the number of trout at t_4 can be greater than the number of trout at t_3, because immigration into the study section can occur between t_3 and t_4 (Section 1.3). Secondly, the error in (6.26) is not distributed $N(0, \sigma_4^2)$, because the errors ϵ_{3i} and ϵ_{4i} are positively correlated (Section 6.4.2). Given that $R_{3i} = 0$, the error ϵ_{3i} must be "large" and negative, so the error ϵ_{4i} will also tend to be large and negative. Hence, if there are no trout in the study section at t_3, the number of trout in the study section at t_4 is likely to be small as well.

Finally, the relationship between E^* and R_4 can be described in a similar way to the relationship between R_4 and R_3. If $R_{4i} \neq 0$, E_i can be eliminated from (6.22) to give

$$E_i^* = \max\{0, R_{4i}\exp(\alpha_5 - \alpha_4) + \epsilon_{45i}\},$$

where

$$\epsilon_{45i} = \epsilon_{5i} - \epsilon_{4i}\exp(\alpha_5 - \alpha_4)$$

However, if $R_{4i} = 0$, E_i can not be eliminated from (6.22) and E_i^* is given by

$$E_i^* = \max\{0, E_i\exp(\alpha_5 - \beta_2 E_i - \zeta_2/(w_i - \xi_2) - \zeta_4'/(w_{i+1} - \xi_4')) + \epsilon_{5i}\}.$$

CHAPTER 7 NONLINEARITY MEASURES FOR PIECEWISE LINEAR CHANGE-POINT MODELS

7.1 INTRODUCTION

Measures that assess the nonlinear behaviour of the nonlinear regression model

$$y_i = f(x_i,\theta) + \epsilon_i, \quad i = 1...n, \tag{7.1}$$

where $\{\epsilon_i\}$ are NID$(0,\sigma^2)$, are described in Chapter 3. Many of these are based on the second derivatives of f with respect to θ evaluated at the least squares estimate $\hat{\theta}$; for example,

i) the Bates and Watts curvature measures,

ii) Box's expression for the bias of $\hat{\theta}$,

iii) Hougaard's expression for the skewness of $\hat{\theta}$,

iv) the asymptotic expression for Lowry and Morton's asymmetry measure.

These are useful measures of nonlinearity, provided that the second derivatives of f do not change markedly in the region of interest about $\hat{\theta}$; if this is not the case, the measures can seriously underestimate the nonlinearity of the model.

For example, consider the nonlinear change-point model

$$y_i = \min(1,\theta x_i) + \epsilon_i, \quad i = 1,2, \tag{7.2}$$

with $x_1 = 1$ and $x_2 = 3$. The solution locus (defined in Section 3.2.1) is given by

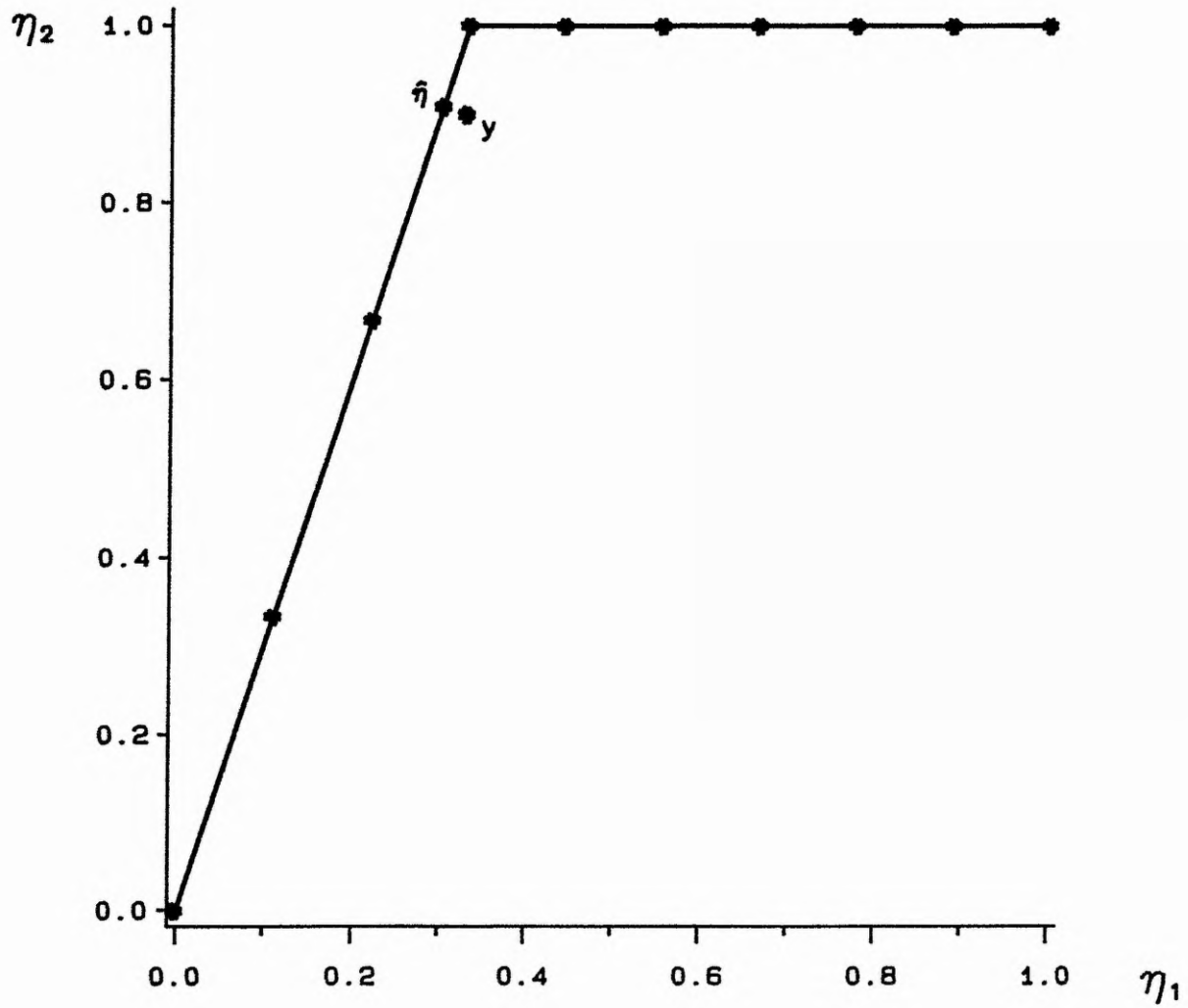


Fig. 7.1 The solution locus of model (7.2).

$$\eta(\theta) = \begin{cases} (\theta,\ 3\theta) & \theta \le 1/3 \\ (\theta,\ 1) & 1/3 < \theta \le 1 \\ (1,\ 1) & 1 < \theta \end{cases}$$

and is shown in Figure 7.1. It is continuous and piecewise linear with the line segments joining at the points corresponding to $\Theta = 1/x_i$, $i = 1,2$. Further, the parameter lines are equispaced on any one line segment, but the spacings vary between line segments.

If $\hat{\eta} = \eta(\hat{\Theta})$ is not at one of the joins, the Bates and Watts curvature measures indicate that a linear approximation to the model about $\hat{\Theta}$ is acceptable in any region of interest, since the second derivatives of f at $\hat{\Theta}$ are zero. However, suppose that $\sigma^2 = 0.01$ and $y_1 = 0.33$, $y_2 = 0.90$, so that $\hat{\Theta} = 0.303$, and suppose that a 95% confidence interval for Θ is required. In this case, a linear approximation is totally inadequate, since the linear approximation interval

$$\theta \in (0.241,\ 0.365)$$

(obtained from equation (3.6)) is much smaller than the exact interval

$$\theta \in (0.241,\ 0.526)$$

(given by the set of values satisfying

$$(y - \eta(\theta))' P(\theta)(y - \eta(\theta)) \le \sigma^2 \chi^2(p;0.05),$$

where $P(\theta)$ is the matrix of projection onto the tangent line at $\eta(\theta)$ (Hamilton, Watts and Bates, 1982)). The Bates and Watts curvature measures seriously underestimate the nonlinear behaviour of the model because some of the derivatives and second derivatives of f are discontinuous in the region of interest about $\hat{\theta}$. (Box's bias, Hougaard's skewness and the asymptotic expression for Lowry and Morton's asymmetry measure also underestimate the nonlinear behaviour of the model for the same reason). Qualitatively, the nonlinearity of model (7.2) depends both on how close $\hat{\eta}$ is to a join and on how much the solution locus "changes" at the join.

Piecewise linear change-point models (ie models similar to (7.2)) are used in the development of a model that describes the relationship between the growth of brown trout fed on maximum rations and time, body weight and water temperature (Chapter 8). Although some study has been made of change-point models, particularly concerning tests for the existence of a change point (eg Hinkley, 1969, 1970; Worsley, 1986; James, James and Siegmund, 1987), no general inference methodology for piecewise linear change-point models is available. Instead, the simplest approach is to make inferences based on a linear approximation. However, the adequacy of the linear approximation can not be assessed by the standard measures of nonlinearity because piecewise linear change-point models have discontinuous derivatives. Hence, in this Chapter, new measures of nonlinearity for piecewise linear change-point models are developed. The principles behind the Bates and Watts curvature measures are used to develop the new measures and are briefly described in Section 7.2. In Section 7.3, the solution locus of a continuous piecewise linear change-point model is shown to be piecewise hyperplanar. In Section 7.4, the new measures are described for a solution locus that consists of just two 2-dimensional planes and in Section

7.5, the case of a multi-hyperplanar solution locus is considered. The measures are applied to a model of the growth rate of brown trout fed on maximum rations in Section 7.6. In Section 7.7, the nonlinear behaviour of alternative parameterisations of piecewise linear change-point models is discussed.

7.2 THE CURVATURE MEASURES OF BATES AND WATTS

Bates and Watts' measures of intrinsic and parameter effects nonlinearity are based on the curvature of the solution locus and of the parameter lines at $\hat{\eta}$. In their notation (Bates and Watts, 1980), an arbitrary straight line in the parameter space through $\hat{\theta}$ can be written as

$$\theta(b) = \hat{\theta} + bh,$$

where h is any non-zero p-vector and where b is a scalar parameter. The image of $\Theta(b)$ in the solution locus is the lifted line

$$\eta_h(b) = \eta(\hat{\theta} + bh). \tag{7.3}$$

The lifted line can be interpreted as the trajectory of a particle whose position is $\eta_h(b)$ at time b.

The tangent vector to $\eta_h(b)$ at $b = 0$ and the second derivative of $\eta_h(b)$ at $b = 0$ are denoted by $\dot{\eta}_h$ and $\ddot{\eta}_h$ respectively, assuming that these derivatives exist. These vectors are the instantaneous velocity and acceleration of the particle at $b = 0$.

The acceleration vector $\ddot{\eta}_h$ can be written as the sum of three components

$$\ddot{\eta}_h = \ddot{\eta}_h^N + \ddot{\eta}_h^G + \ddot{\eta}_h^P. \tag{7.4}$$

where

i) the normal acceleration $\ddot{\eta}_h^N$ is normal to the tangent plane and determines the change of direction of $\dot{\eta}_h$ normal to the tangent plane,

ii) the geodesic acceleration $\ddot{\eta}_h^G$ is parallel to the tangent plane but normal to $\dot{\eta}_h$ and determines the change of direction of $\dot{\eta}_h$ in the tangent plane,

iii) the parallel acceleration $\ddot{\eta}_h^P$ is parallel to $\dot{\eta}_h$ and determines the change in speed of the moving point.

The tangential acceleration can be written as one component

$$\ddot{\eta}_h^T = \ddot{\eta}_h^G + \ddot{\eta}_h^P. \tag{7.5}$$

The normal and tangential curvatures in direction $\dot{\eta}_h$ are given by

$$K_h^N = \left\| \ddot{\eta}_h^N \right\| / \left| \dot{\eta}_h \right|^2 \tag{7.6}$$

$$K_h^T = \left\| \ddot{\eta}_h^T \right\| / \left| \dot{\eta}_h \right|^2 \tag{7.7}$$

respectively. These curvatures can be identified with the intrinsic and parameter effects nonlinearity of the model respectively and are known as the intrinsic and

parameter effects curvatures in direction $\dot{\eta}_h$. If s^2 is an estimate of σ^2 based on ν degrees of freedom and if $\rho = s\sqrt{p}$ is the standard radius, the relative intrinsic curvature in direction $\dot{\eta}_h$ is

$$\gamma_h^N = K_h^N \rho \tag{7.8}$$

and the relative parameter-effects curvature in direction $\dot{\eta}_h$ is

$$\gamma_h^T = K_h^T \rho. \tag{7.9}$$

A linear approximation to a nonlinear model involves two distinct approximations: a tangent plane approximation and a uniform co-ordinates approximation (Section 3.2.2). If $\gamma_h{}^N$ is small compared to $1/\sqrt{F(p,\nu;\alpha)}$ (ie less than $1/(2\sqrt{F})$) for all vectors h, the tangent plane approximation is likely to be good in the $100(1-\alpha)\%$ confidence region of interest and inferences based on the tangent plane approximation can be made with security. Similarly, if $\gamma_h{}^T$ is small compared to $1/\sqrt{F(p,\nu;\alpha)}$ for all h, the uniform co-ordinates approximation is likely to be good in the region of interest.

The curvature measures $K_h{}^N$ and $K_h{}^T$ compare the linear approximation in direction $\dot{\eta}_h$ with circular approximations. For example, $K_h{}^N$ is the inverse of the radius of the circle that best approximates the solution locus in direction $\dot{\eta}_h$.

7.3 CONTINUOUS PIECEWISE LINEAR CHANGE-POINT MODELS

Consider the continuous piecewise linear model with unknown change points

$$y_i = f_j(x_i, \theta) + \epsilon_i \quad \text{if } x_i \in \chi_j(\theta), \quad j = 1...m_1, \; i = 1...n, \tag{7.10}$$

where

i) the functions $f_j(x, \theta)$ are linear in Θ, $j = 1...m_1$,

ii) χ, the set of all possible design points x, is the union of the disjoint sets $\chi_j(\Theta)$, $j = 1...m_1$,

iii) the model function

$$f(x, \theta) = \sum_{j=1}^{m_1} f_j(x, \theta) I(x \in \chi_j(\theta))$$

is a continuous function of x and Θ, where I(.) is the indicator function of the event in brackets. For example, model (7.2) can be written as

$$y_i = \begin{cases} \theta x_i + \epsilon_i & \text{if } x_i \leq 1/\theta. \\ 1 + \epsilon_i & \text{if } x_i > 1/\theta. \end{cases}$$

In this case, $m_1 = 2$, and

$$f_1(x, \theta) = \theta x,$$

$$f_2(x, \theta) = 1,$$

$$\chi_1(\theta) = \{x; x \leq 1/\theta\}.$$

$$\chi_2(\theta) = \{x; x > 1/\theta\}.$$

Since f_j is linear in Θ, $j = 1...m_1$, the solution locus can be written as

$$\eta(\theta) = \zeta_j + X_j\theta \quad \text{if} \quad \theta \in \Theta_j, \quad j = 1...m_2, \tag{7.11}$$

where

i) ζ_j are offset n-vectors, $j = 1...m_2$,

ii) X_j are nxp matrices, $j = 1...m_2$, (where p is the dimension of the parameter space; ie $\Theta = (\Theta_1,...,\Theta_p)'$),

iii) the parameter space Θ is the union of the disjoint sets Θ_j, $j = 1...m_2$.

In the case of model (7.2), $m_2 = 3$ and

$$\Theta_1 = \{\theta; \theta \leq 1/3\},$$

$$\Theta_2 = \{\theta; 1/3 < \theta \leq 1\},$$

$$\Theta_3 = \{\theta; 1 < \theta\},$$

$$\zeta_1 = (0,0)', \quad X_1 = (1,3)',$$

$$\zeta_2 = (0,1)', \quad X_2 = (1,0)',$$

$$\zeta_3 = (1,1)', \quad X_3 = (0,0)',$$

The solution locus (7.11) is piecewise hyperplanar and the parameter lines on any segment of the solution locus are straight, parallel and equispaced. In addition, the solution locus is continuous, since the model function $f(x,\Theta)$ is continuous in Θ. These special properties enable the solution locus and the parameter lines to be

compared to the linear approximation tangent plane and the "uniform co-ordinate lines" in a comparatively large region about $\hat{\eta}$ and hence enable the nonlinear behaviour of the model to be assessed.

Some notation is introduced first. Let H_j be the hyperplane

$$H_j = \{\zeta_j + X_j\theta; \theta \in R^p\}.$$

The solution locus segment (sl-segment) of H_j is the subset

$$sl(H_j) = \{\zeta_j + X_j\theta; \theta \in \Theta_j\}$$

that forms part of the solution locus. In model (7.2)

$$\left\{\begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\theta; 1/3 < \theta \leq 1\right\}$$

is the sl-segment of the line

$$\left\{\begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\theta; \theta \in R\right\}$$

The matrix X_j is called the matrix associated with H_j and the rank space of X_j is

$$rank(X_j) = \{X_j\theta; \theta \in R^p\}.$$

The orthogonal projection of a vector z onto hyperplane H is denoted by $P_H z$; all projections mentioned in the text are orthogonal projections.

7.4 SOLUTION LOCUS OF TWO PLANES

It is convenient to consider first the case where the solution locus consists of the sl-segments of two 2-dimensional planes A and B that intersect in a 1-dimensional line C in n-dimensional sample space. The solution locus is 3-dimensional, so it can be represented graphically (Fig. 7.2) (although it is important to remember that it is embedded in an n-dimensional space). In Figure 7.2, sl(A) and sl(B) are denoted by the solid lines. Assume that $\hat{\eta}$ lies on the interior of sl(A).

7.4.1 Half Lifted Lines

Consider the lifted line (7.3)

$$\eta_A(b) = \eta(\hat{\theta} + bh).$$

Since the parameter lines are straight, parallel and equi-spaced on both sl-segments and are continuous everywhere, the lifted line can take one of two forms:

i) $\eta_h(b)$ is a straight line parallel to C, that lies entirely on sl(A) and "travels at constant speed" (that is, the speed of the particle at $\eta_h(b)$ at time b is constant) (Fig. 7.3a),

ii) $\eta_h(b)$ consists of two joined line segments, one on sl(A) and the other on sl(B); both line segments are straight and travel at constant speed (although in general, the two directions and speeds are different) (Fig. 7.3b).

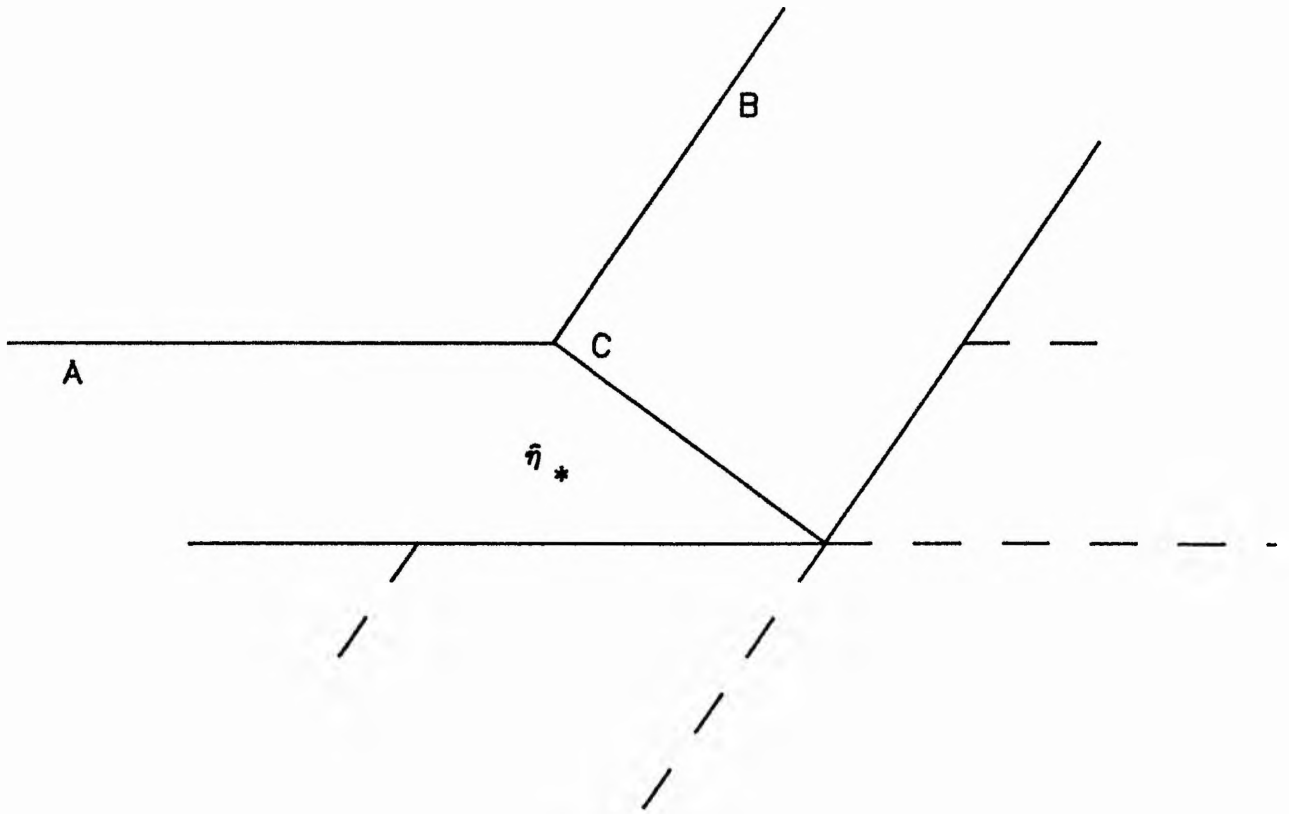


Fig. 7.2 The solution locus consisting of two planes.

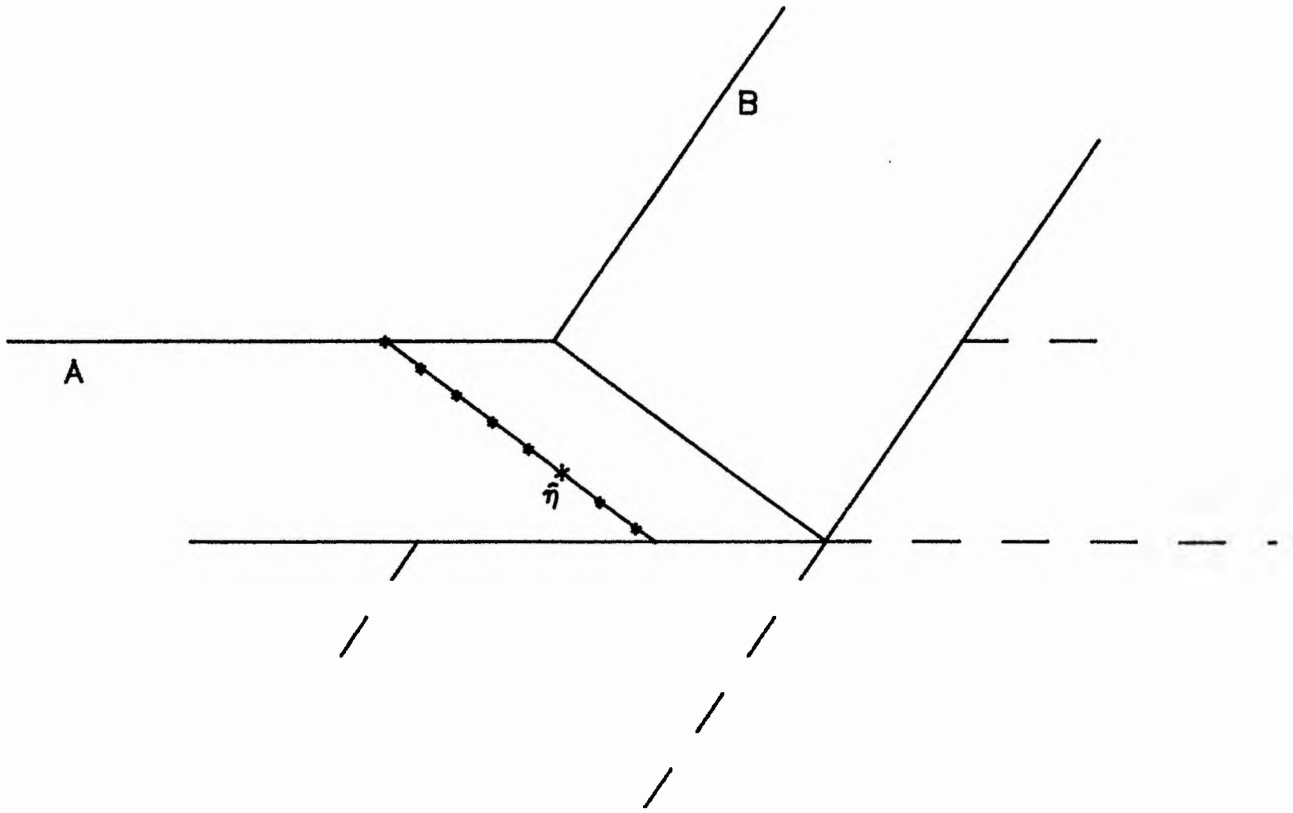


Fig. 7.3a Lifted lines.

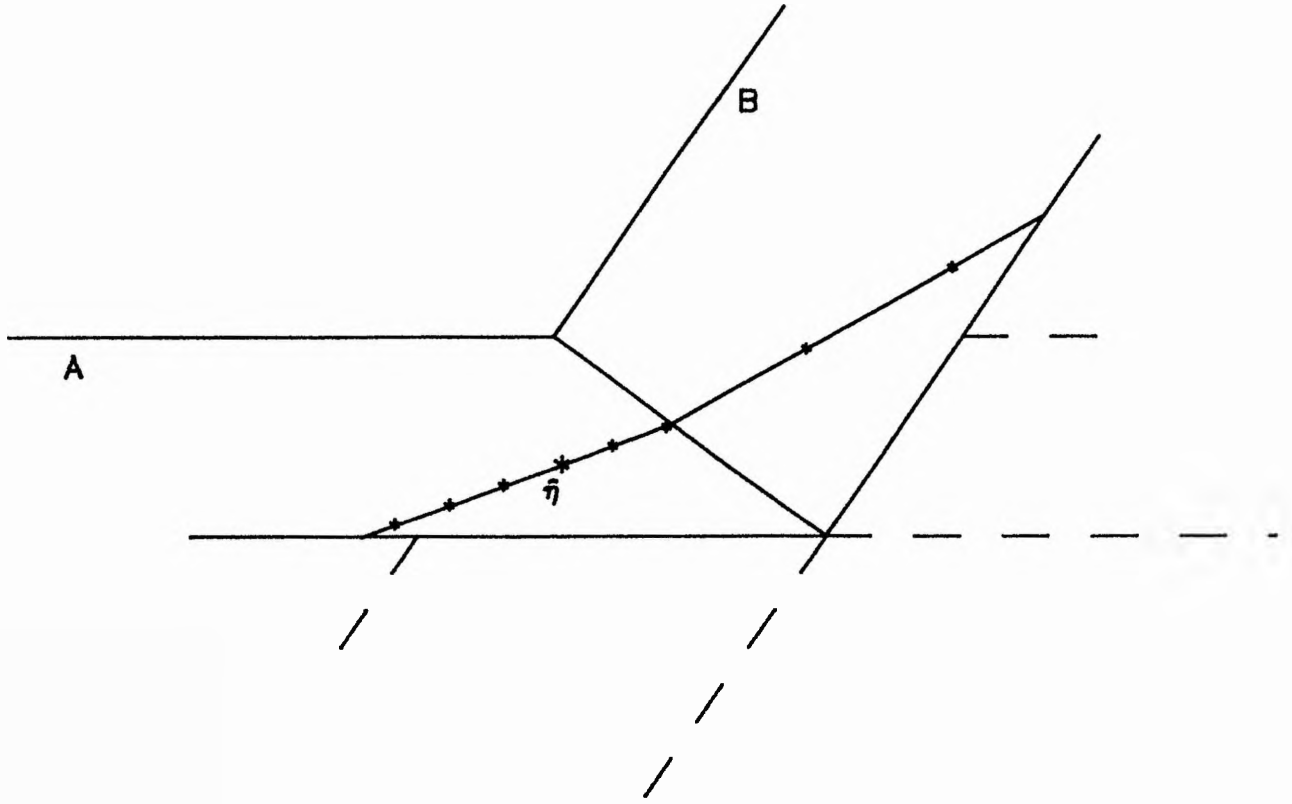


Fig. 7.3b Lifted lines.

In case ii) above, the two "halves" of the lifted line corresponding to positive and negative values of b behave quite differently; one half lies entirely on sl(A), the other lies on both sl(A) and sl(B). Thus, it is convenient to consider these halves separately. Define the half lifted line $\eta_h^+(b)$ as the portion of the lifted line $\eta_h(b)$ corresponding to values of $b \geq 0$. This can be viewed as the trajectory of a particle travelling outwards from $\hat{\eta}$. Either the particle always travels on sl(A) (Fig. 7.4a) or it travels on sl(A) until it "hits" B and then travels on sl(B) (Fig. 7.4b). The behaviour of the original lifted line $\eta_h(b)$ is investigated by considering the two half lifted lines $\eta_h^+(b)$ and $\eta_{-h}^+(b)$.

7.4.2 Piecewise Linear Curvatures

By decomposing the acceleration vector $\ddot{\eta}_h$ into three components (see equation (7.4)), Bates and Watts identify three types of nonlinearity associated with the lifted line $\eta_h(b)$:

i) the effect of the change of direction of $\dot{\eta}_h$ normal to the tangent plane,

ii) the effect of the change of direction of $\dot{\eta}_h$ parallel to the tangent plane,

iii) the effect of the change of speed of the moving point.

However, by only using the derivatives of $\eta_h(b)$ at $b = 0$, Bates and Watts only consider the behaviour of $\eta_h(b)$ in a small neighbourhood about $\hat{\eta}$.

In the present case, the special form of the half lifted line $\eta_h^+(b)$ enables the nonlinear behaviour associated with all of $\eta_h^+(b)$ to be represented by three effects very similar to those identified by Bates and Watts. If $\eta_h^+(b)$ lies entirely on sl(A), there is no nonlinear behaviour associated with $\eta_h^+(b)$. If $\eta_h^+(b)$ lies on both sl(A) and sl(B), it can be represented by the line segments RS and ST as in Figure 7.5;

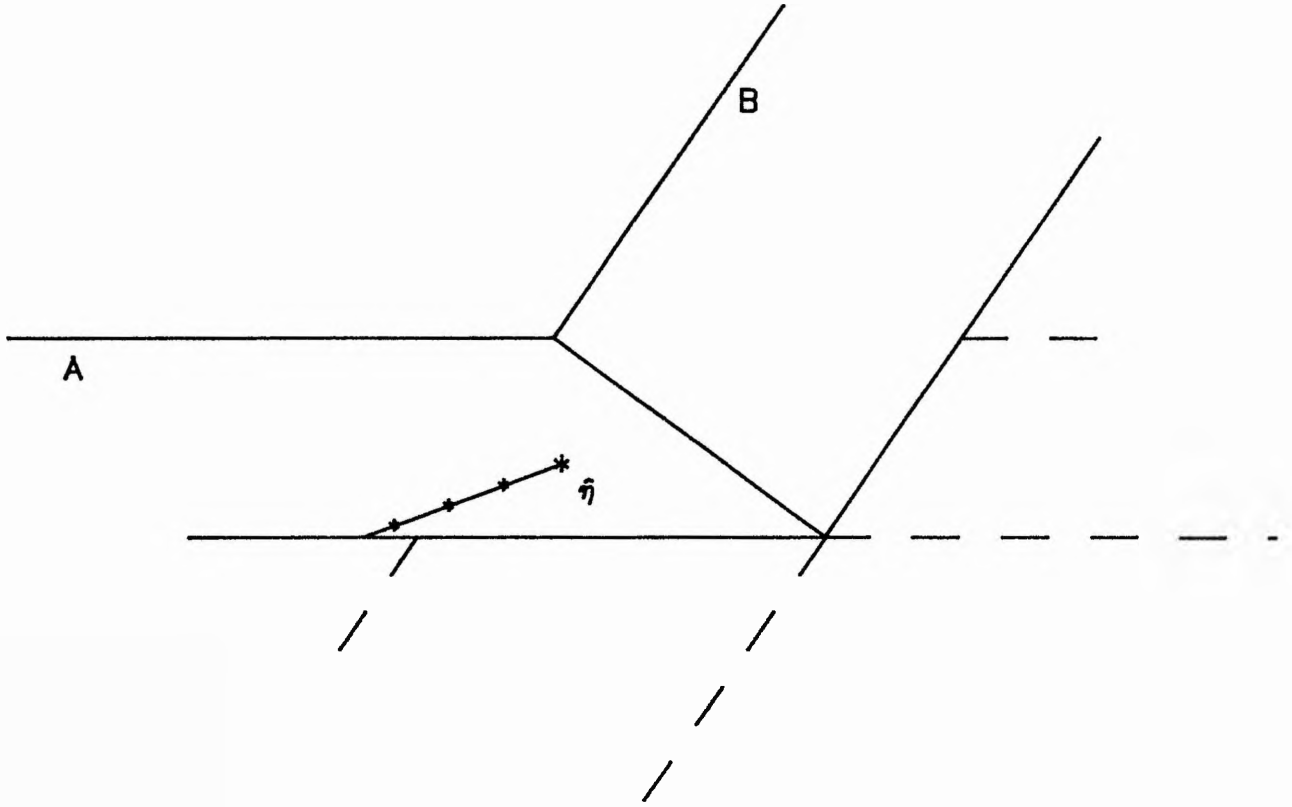


Fig. 7.4a Half lifted lines.

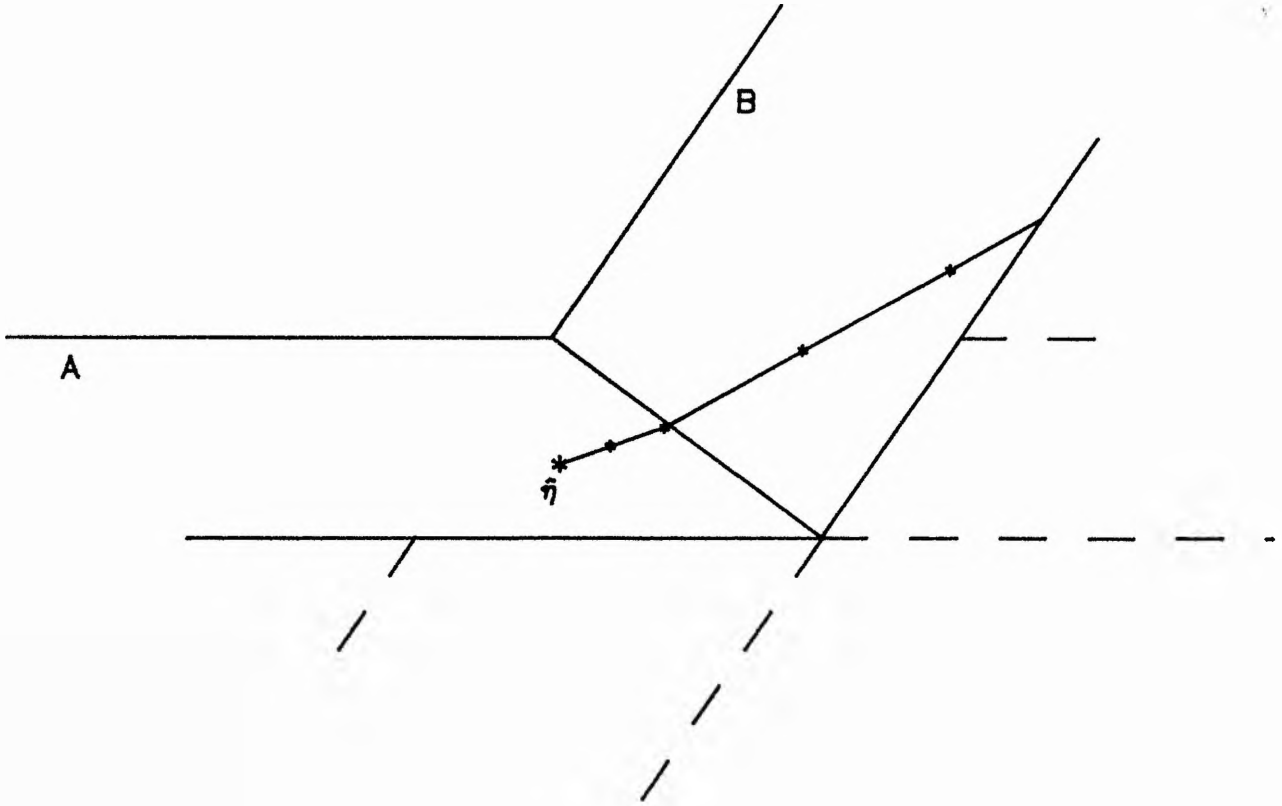


Fig. 7.4b Half lifted lines.

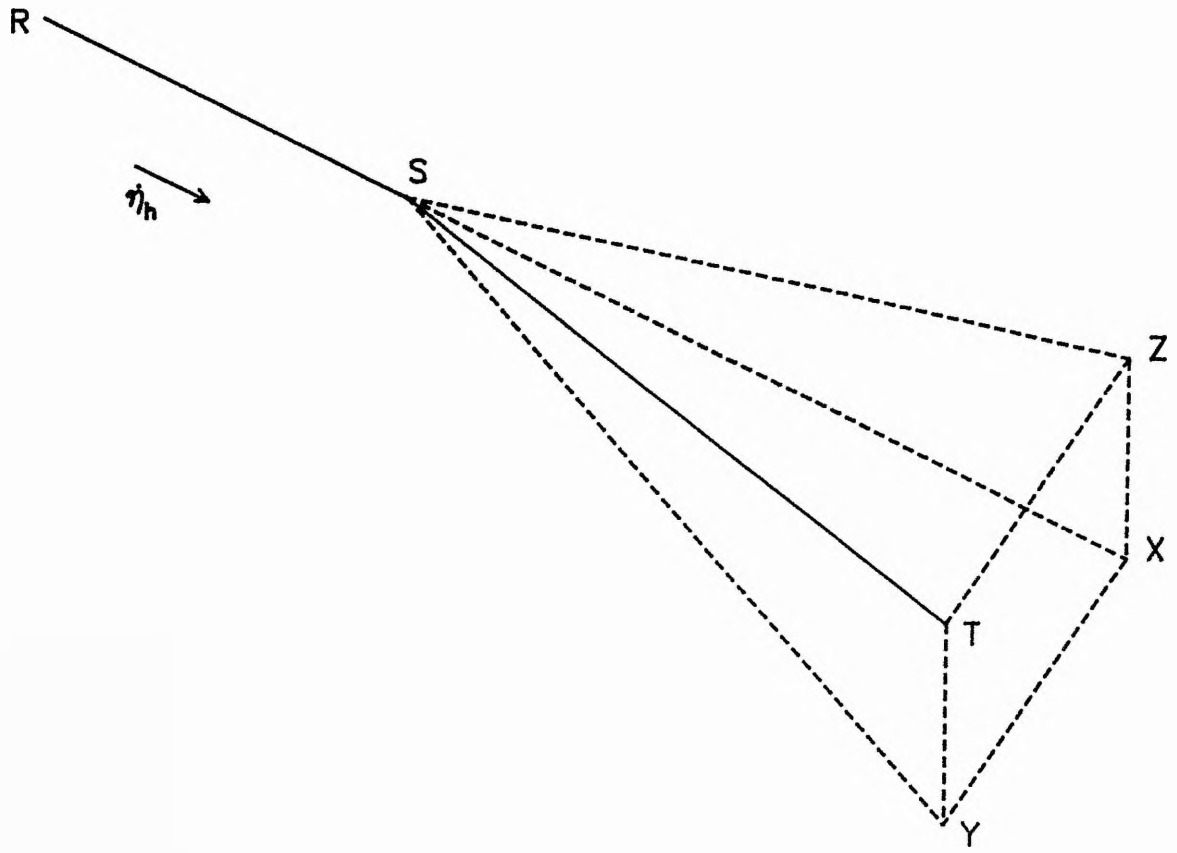


Fig. 7.5 The half lifted line $\eta_h^+(b)$.

R corresponds to $\hat{\eta}$ and S is the point at which $\eta_h^+(b)$ hits B, so that the tangent vector $\dot{\eta}_h$ is in the direction RS. Let V_h be the plane (in the 3-dimensional hyperplane containing A and B) orthogonal to A and containing RS. Let

i) SZ be the projection of ST onto V_h,

ii) SY be the projection of ST onto A,

iii) RSX be a straight line (so that SX is the projection of SZ onto A and also the projection of SY onto V_h).

Thus, the points $RSXY$ lie in A, STZ lie in B and $RSZX$ lie in V. The nonlinear behaviour of $\eta_h^+(b)$ is decomposed into the three components

i) the normal component which measures the effect of the change of direction between the line segments RS and SZ (ie, normal to the tangent plane A),

ii) the geodesic component which measures the effect of the change of direction between the line segments RS and SY (ie, parallel to A, but normal to $\dot{\eta}_h$),

iii) the parallel component which measures the effect of the change in speed in the line RSX (ie, parallel to $\dot{\eta}_h$).

The segments RS and SZ lie in the solution locus. Regardless of the parameterisation of the model, they are defined uniquely by $\dot{\eta}_h$ as the section of the solution locus in direction $\dot{\eta}_h$. Thus, the normal component can be regarded as the intrinsic component of nonlinearity. The geodesic and parallel components combine to give the parameter effects component of nonlinearity.

The effect of the changes of direction and speed depends on both the magnitudes of the changes and on the distance from $\hat{\eta}$ to the point at which $\hat{\eta}_h^+(b)$ hits B (ie $\|RS\|$). For example, if the region of interest about $\hat{\eta}$ is small compared to $\|RS\|$,

the behaviour of $\eta_h^+(b)$ on B has a negligible effect on the nonlinearity of the model. Also, if the region of interest is large compared to $\|RS\|$ but $\eta_h^+(b)$ hardly deviates when it changes from sl(A) to sl(B), the nonlinear behaviour of $\eta_h^+(b)$ will be small.

The nonlinear behaviour of $\eta_h^+(b)$ is measured by approximating the instantaneous changes of direction and speed at S by smooth changes of direction and speed that take effect from the time the particle leaves $\hat{\eta}$ (as described below). These smooth changes are then converted to curvatures, known as piecewise linear (or pl-) curvatures.

7.4.2.1 Intrinsic Pl-curvature

The projection of $\eta_h^+(b)$ onto V_h consists of two line segments RS and SZ, lying in A and B respectively, with an angle $w_h^N \in (0, \pi/2)$ between them (Fig. 7.6). The Bates and Watts intrinsic curvature in direction $\dot{\eta}_h$ (equation (7.6)) is the inverse of the radius of the circle that best approximates the solution locus in direction $\dot{\eta}_h$ in a small neighbourhood about $\hat{\eta}$. An analogous approach here is to approximate the line segments RS and SZ by a circle that touches both line segments and that touches RS at R (ie at $\hat{\eta}$). The intrinsic pl-curvature c_h^N is defined as the inverse of the radius of this circle and is given by

$$\begin{aligned} c_h^N &= \frac{\cot((\pi - w_h^N)/2)}{\xi_h} \\ &= \frac{\tan(w_h^N/2)}{\xi_h} \\ &= \frac{\sin(w_h^N)}{\xi_h(1 + \cos(w_h^N))} \end{aligned} \qquad (7.12)$$

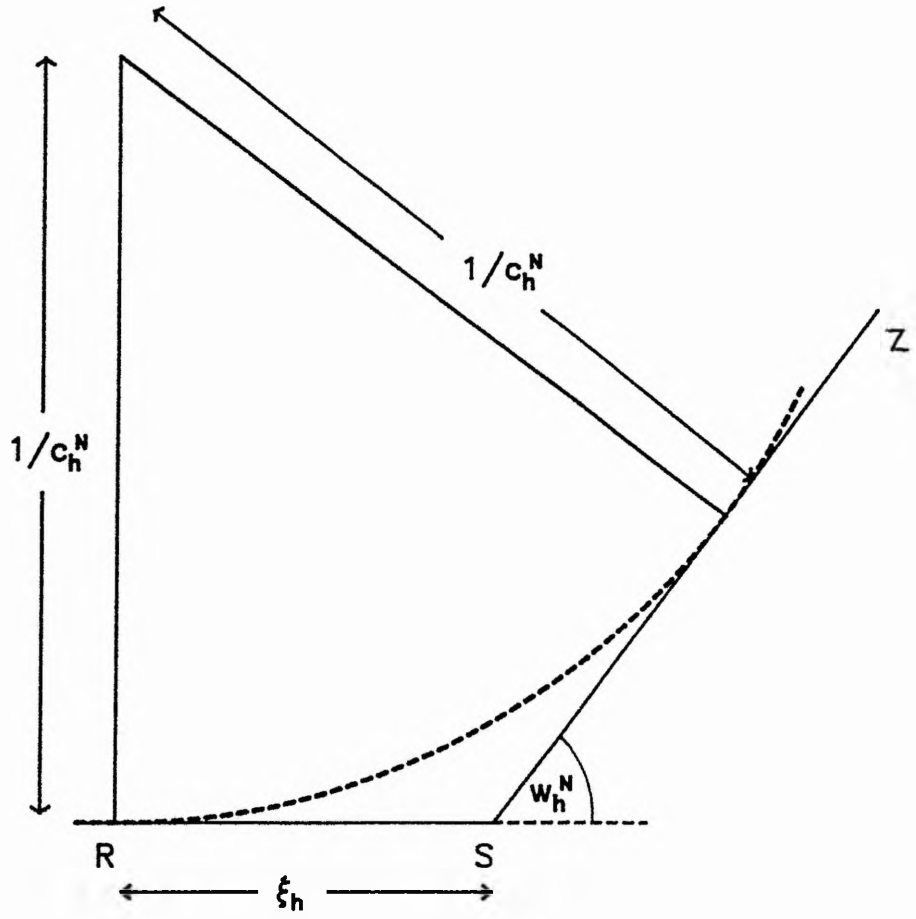


Fig. 7.6 The circular approximation to the solution locus.

where $\xi_h = \|RS\|$. The relative intrinsic pl-curvature in direction $\dot{\eta}_h$ is

$$\Lambda_h^N = \rho c_h^N, \tag{7.13}$$

where ρ is the standard radius (cf equation (7.8)). (If $\eta_h^+(b)$ lies entirely on sl(A) then $\Lambda_h{}^N = 0$).

7.4.2.2 Parameter effects Pl-curvature

The projection of $\eta_h^+(b)$ onto A consists of two line segments RS and SY with an angle $w_h{}^G \in [0, \pi/2]$ between them. By a similar argument to before, the geodesic pl-curvature in direction $\dot{\eta}_h$ is defined as the inverse of the radius of the circle that touches both segments and also touches RS at $\hat{\eta}$ and is given by

$$c_h^G = \frac{\sin(w_h^G)}{\xi_h(1 + \cos(w_h^G))} \tag{7.14}$$

The projection of $\eta_h^+(b)$ onto $A \cap V_h$ is the straight line RSX; the "speed" of the line is $\|\dot{\eta}_h\|$ on RS and v_h say on SX. The change of speed at S can be represented by a pair of line segments with an angle

$$w_h^P = \tan^{-1}\left(\frac{|v_h - |\dot{\eta}_h||}{|\dot{\eta}_h|}\right)$$

between them (Fig. 7.7). The parallel pl-curvature in direction $\dot{\eta}_h$ is defined as the inverse of the radius of the circle that approximates these line segments and is given by

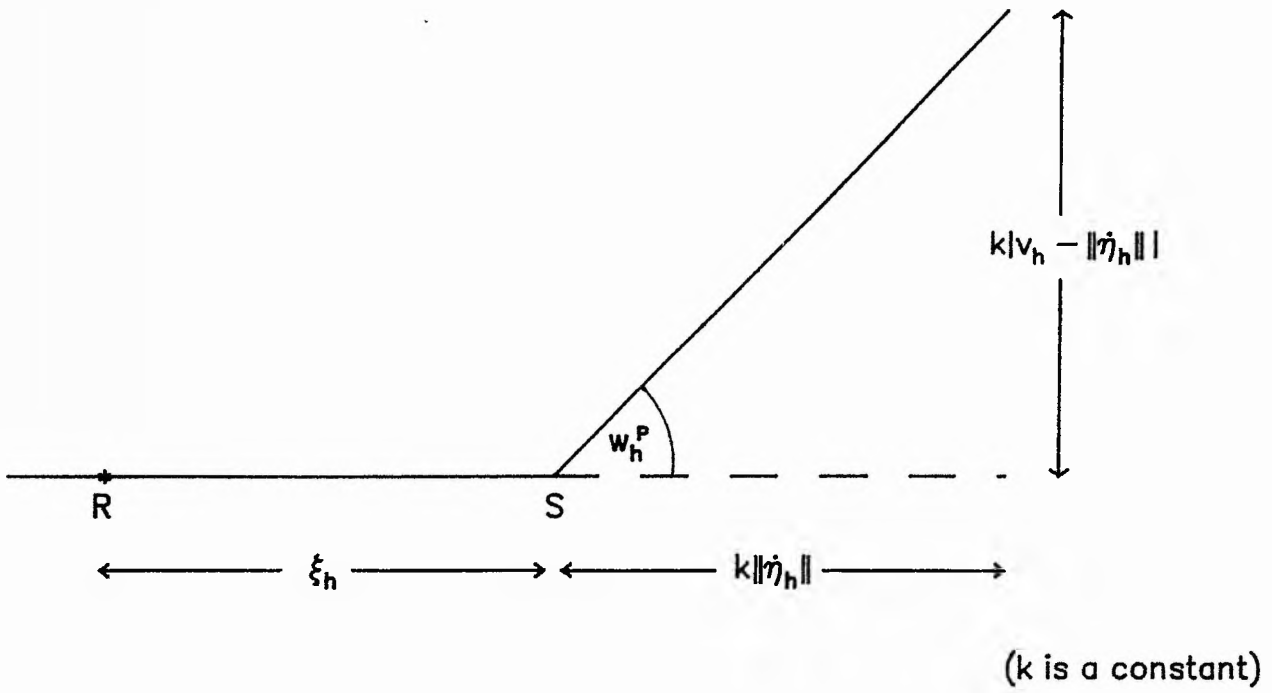


Fig. 7.7 Line segments representing the change of speed at S.

$$c_h^r = \frac{\sin(w_h^r)}{\xi_h(1 + \cos(w_h^r))}$$

$$= \frac{|v_h - |\dot{\eta}_h||}{\xi_h(|\dot{\eta}_h| + \sqrt{(v_h^2 - 2v_h|\dot{\eta}_h| + 2|\dot{\eta}_h|^2)})} \qquad (7.15)$$

The parameter effects pl-curvature in direction $\dot{\eta}_h$ is given by

$$c_h^T = \sqrt{((c_h^c)^2 + (c_h^r)^2)},$$

(cf (7.5), (7.7)) and the relative parameter effects pl-curvature in direction $\dot{\eta}_h$ is

$$\Lambda_h^T = \rho c_h^T. \qquad (7.16)$$

(cf (7.9)).

7.4.2.3 Conservative measures

The pl-curvatures are conservative measures of nonlinearity because the circular approximations "lie above" the line segments they approximate. For example, the circular approximation to the line segments *RS* and *SZ* (used to calculate Λ_h^N) is more "nonlinear" than the line segments *RS* and *SZ* in the sense that it deviates more from the tangent plane *A* (Fig. 7.6). Thus, if Λ_h^N is small compared to $1/\sqrt{F(2,\nu;\alpha)}$ for all *h*, the tangent plane approximation is good in the region of interest. If Λ_h^N is large for some *h*, say h_o, the tangent plane approximation is likely to be poor (although it is possible that $\Lambda_{h_o}^N$ is large only because ξ_{h_o} is very small; this situation is discussed in Section 7.8). Similarly, if Λ_h^T is small for all *h*, the uniform co-ordinates approximation is good in the region of interest.

7.4.3 Linear Transformations of the Parameter Space

The calculation of the pl-curvatures is simplified by considering the effect of a linear transformation of the parameter space co-ordinates, similar to that used by Bates and Watts (1980). The θ parameter lines on sl(A) are straight, parallel and equispaced (Section 7.3). Hence, the parameter lines of any linear transformation of the θ co-ordinates are also straight, parallel and equispaced. In particular, there exists a linear transformation $\theta \mapsto \phi$ such that, in addition,

i) the ϕ parameter lines on sl(A) are orthogonal,

ii) the unit circle in the ϕ parameter space maps to the unit circle of centre $\hat{\eta}$ lying in A (assuming, for one moment, that this circle lies completely within sl(A)), (Fig. 7.8).

The half lifted lines are the same for both the θ and the ϕ parameterisations, since the transformation $\theta \mapsto \phi$ is linear. Also, Λ_h^N and Λ_h^T are independent of the length of h, so it is sufficient to consider the half lifted lines corresponding to unit vectors in the ϕ parameterisation to obtain the curvatures in every possible direction (cf Bates and Watts 1980). By property ii) above, this is equivalent to considering the half lifted lines that "head" in direction d for all unit vectors d in rank(X), where

i) X is the matrix associated with A (Section 7.3),

ii) the half lifted line $\eta_h^+(b)$ heads in direction d if $\eta_h^+(b) = \hat{\eta} + bd$ for $b \geq 0$ and b sufficiently small (Fig. 7.9).

Denote the half lifted line (that heads) in direction d by $\eta_d^+(b)$. To avoid confusion, all structures with a suffix h pertain to the half lifted line

$$\eta_h^+(b) = \eta(\hat{\theta} + bh)$$

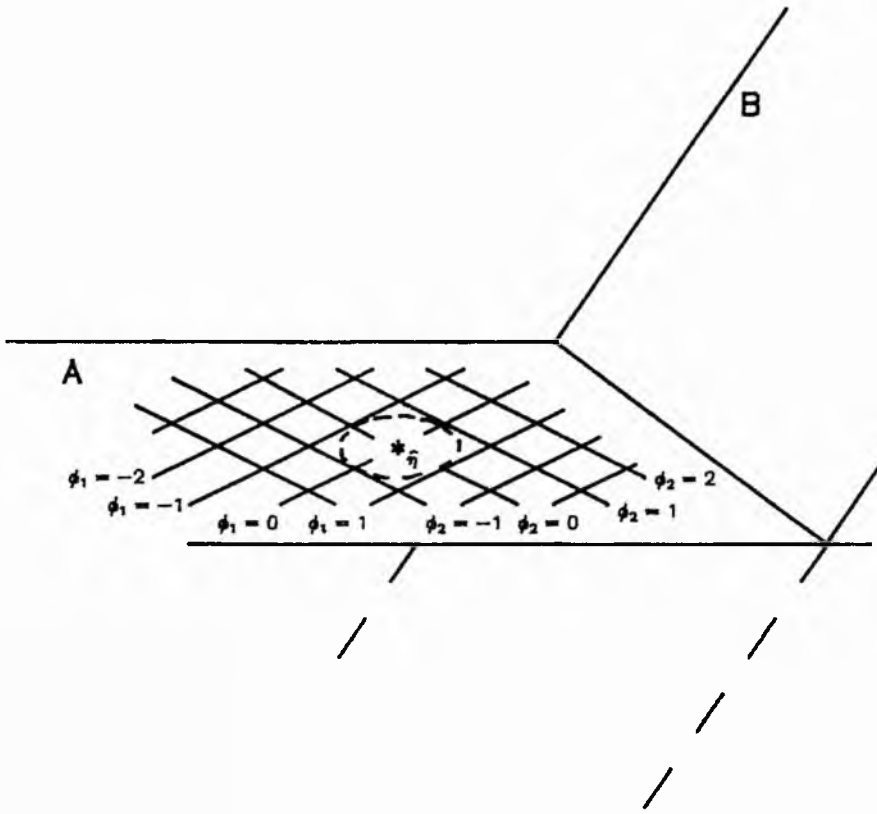
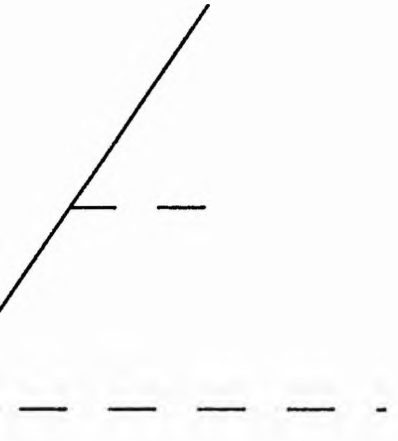


Fig. 7.8 The ϕ parameter lines.



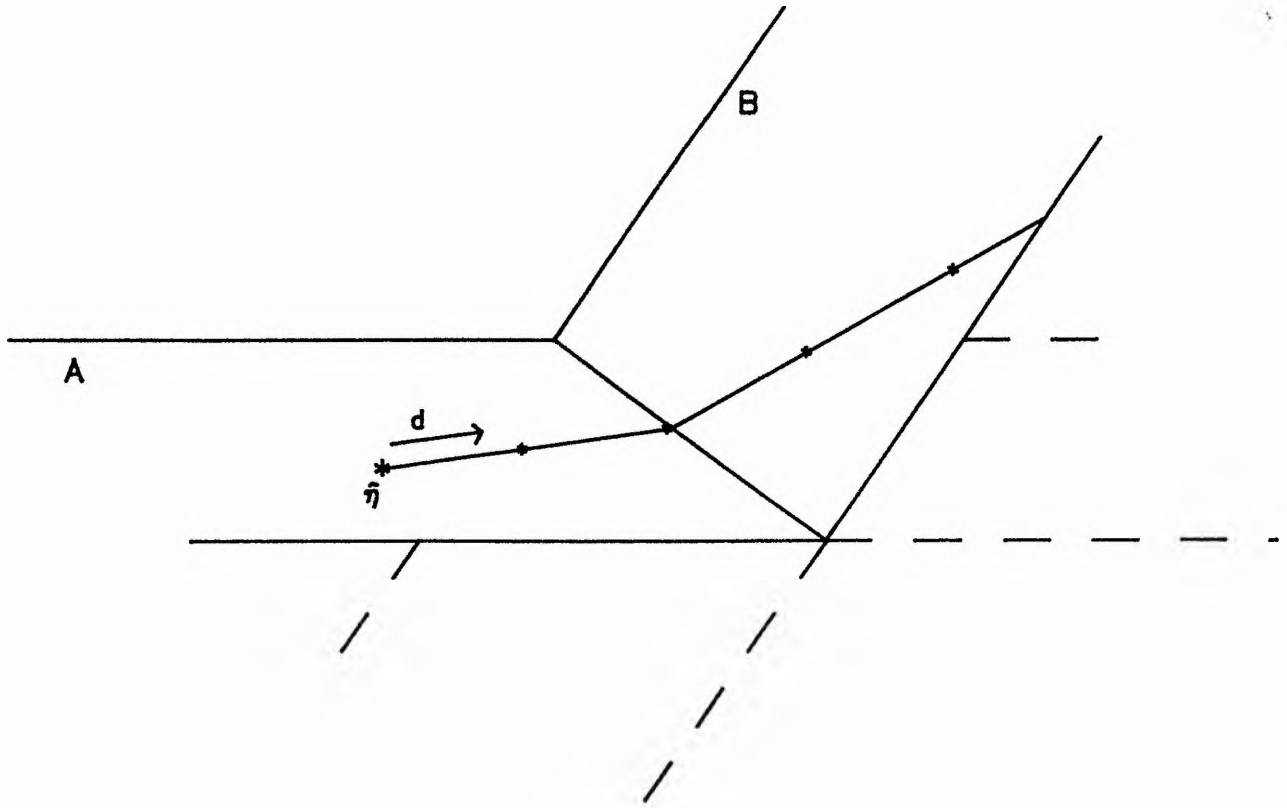


Fig. 7.9 The half lifted line in direction d.

in the original Θ co-ordinates. Any structure with any other vector suffix, say d, pertains to the half lifted line in the direction of that vector.

7.4.4 Calculating Pl-Curvatures

The calculation of the pl-curvatures requires the construction of a system of vectors which can be used to describe the relationship between the planes A and B and the parameter lines. In this two-planar case, the notation is quite cumbersome; however, the multi-hyperplanar case is easier to follow if the notation is introduced here. Let Y be the matrix associated with B. Let e_2 be a unit vector in the direction of the line of intersection C and let e_1 and e_B be unit vectors orthogonal to e_2 lying in rank (X) and rank (Y) respectively (Fig. 7.10); the directions of e_1 and e_B are chosen so that $\eta_{|e_1}{}^{+}(b)$ hits C and $e_1{}' e_B > 0$. Then $\{e_1,e_2\}$ is an orthonormal basis of rank (X) and $\{e_B,e_2\}$ is an orthonormal basis of rank (Y).

Given $\{e_1,e_2\}$, the unit vectors d in rank (X) can be parameterised by an angle $\tau \in (-\pi,\pi\,]$ as

$$\begin{aligned} d \;=\; d(\tau) \;&=\; \cos(\tau)e_1 \;+\; \sin(\tau)e_2 \\ &=\; [e_1,e_2]g, \end{aligned}$$

where

$$g(\tau) \;=\; \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \;=\; \begin{pmatrix} \cos(\tau) \\ \sin(\tau) \end{pmatrix}$$

The pl-curvatures of $\eta_d{}^{+}(b)$ can then be expressed in terms of e_1, e_B, e_2 and τ.

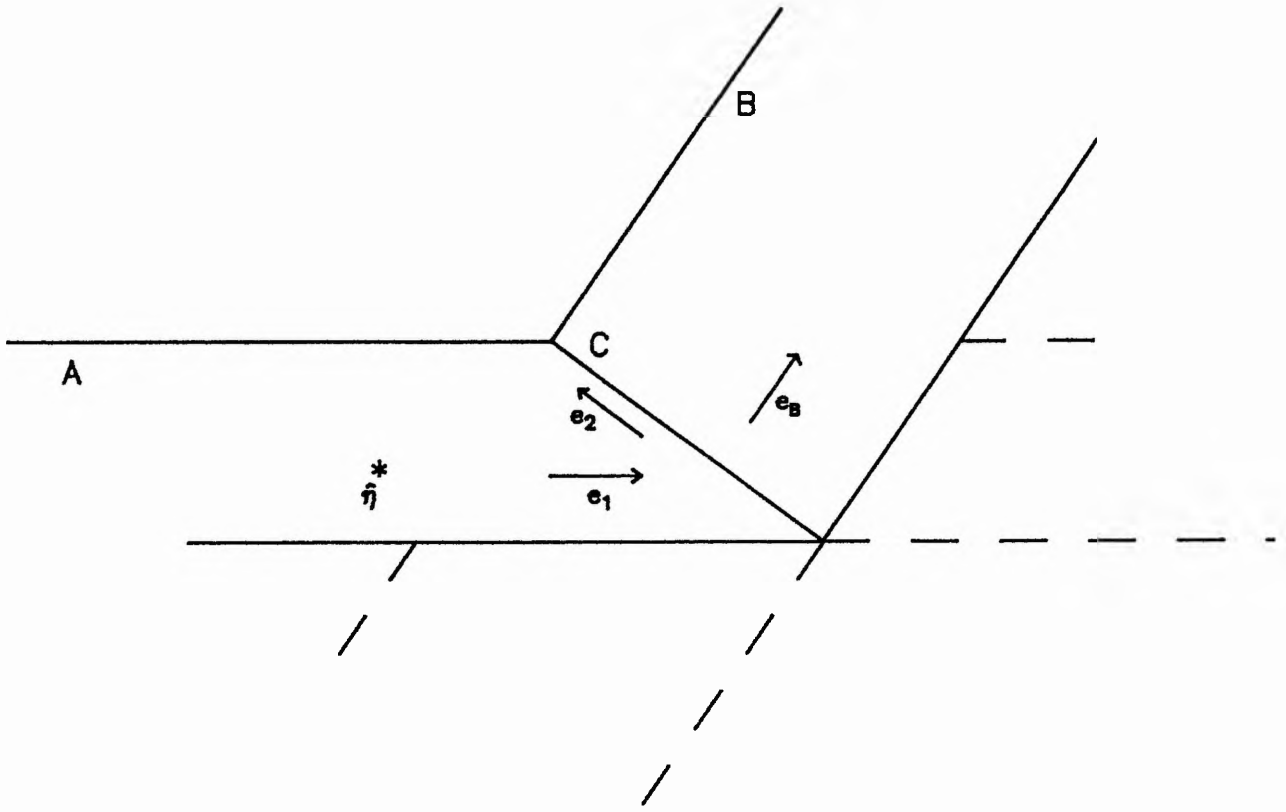


Fig. 7.10 The unit vectors e_1, e_2 and e_B.

Let $w^N \in (0, \pi/2]$ be the angle between A and B and let ξ be the shortest distance from $\hat{\eta}$ to C. These correspond to the angle between the solution locus segments in direction e_1 and the distance from $\hat{\eta}$ to C in direction e_1 respectively, so that $w^N = w_{e_1}{}^N$ and $\xi = \xi_{e_1}$. The projection of e_B onto rank (X) is then

$$P_{rank(X)} e_B = \cos(w^N) e_1 .$$

7.4.4.1 Intrinsic Pl-curvature

If $\tau \in (-\pi, -\pi/2] \cup [\pi/2, \pi]$, then $\eta_d{}^+(b)$ lies entirely on $sl(A)$ and $\Lambda_d{}^N = 0$. Otherwise, the projection of $\eta_d{}^+(b)$ onto V_d consists of two line segments, one in direction d and one in direction d^N, where d^N is a unit vector given by

$$d^N = [e_B, e_2] g^N ,$$

for some 2-vector $g^N = (g_1{}^N, g_2{}^N)'$. Since V_d is orthogonal to A and parallel to d, the projection of the line segments onto A is in direction d. Thus, the projection of d^N onto rank (X),

$$P_{rank(X)} d^N = g_1^N \cos(w^N) e_1 + g_2^N e_2 ,$$

is parallel to d. Hence

$$g^N = \frac{1}{\delta} \left(\frac{g_1}{\cos(w^N)}, \; g_2 \right)' ,$$

where δ is a normalising constant given by

$$\delta^2 = \frac{\cos^2(\tau)}{\cos^2(w^N)} + \sin^2(\tau)$$

$$= \frac{1-\sin^2(\tau)\ \sin^2(w^N)}{\cos^2(w^N)}$$

The angle w_d^N between the two line segments is given by

$$\cos(w_d^N) = d'd^N = 1/\delta$$

and the distance from $\hat{\eta}$ to C in direction d is

$$\xi_d = \xi/\cos(\tau).$$

Hence, using (7.12) and (7.13), the relative intrinsic pl-curvature in direction d is

$$\Lambda_d^N = \frac{\rho\ \cos(\tau)\sqrt{(\delta^2 - 1)}}{\xi(\delta + 1)}$$

$$= \frac{\rho\ \cos^2(\tau)\ \sin(w^N)}{\xi(\cos(w^N) + \sqrt{\{1-\sin^2(\tau)\ \sin^2(w^N)\}})} \qquad (7.17)$$

To calculate the maximum intrinsic pl-curvature $\Lambda^N{}_{max}$ and the RMS intrinsic pl-curvature $\Lambda^N{}_{RMS}$ it is sufficient to know e_1, e_B and hence w^N and ξ explicitly. These are found by calculating the projections

$$\eta' = P_B\hat{\eta},$$

$$\eta'' = P_A\eta',$$

$$\eta''' = P_B\eta'',$$

(Fig. 7.11). Then

$$e_1 = (\eta'' - \hat{\eta})/\|\eta'' - \hat{\eta}\|,$$

$$e_B = (\eta''' - \eta')/|\eta''' - \eta'|,$$

and hence,

$$\cos(w^N) = e_1' e_B,$$

$$\xi = |\eta' - \hat{\eta}|/\sin(w^N).$$

The RMS intrinsic pl-curvature is given by

$$(\Lambda^N_{RMS})^2 = \frac{1}{S(2)} \int_{|d|=1} (\Lambda^N_d)^2 dS \tag{7.18}$$

where $S(p)$ is the surface area of the unit sphere in p-dimensional space

$$S(p) = 2(\pi)^{p/2}/\Gamma(p/2). \tag{7.19}$$

Equation (7.18) is expressed in terms of τ as

$$(\Lambda^N_{RMS})^2 = \frac{1}{S(2)} \int_{-\pi}^{\pi} (\Lambda^N_d)^2 d\tau$$

$$= \frac{1}{S(2)} \int_{-\pi/2}^{\pi/2} \frac{\rho \cos^2(\tau) \sin(w^N) d\tau}{\xi(\cos(w^N) + \sqrt{\{1-\sin^2(\tau) \sin^2(w^N)\}})}$$

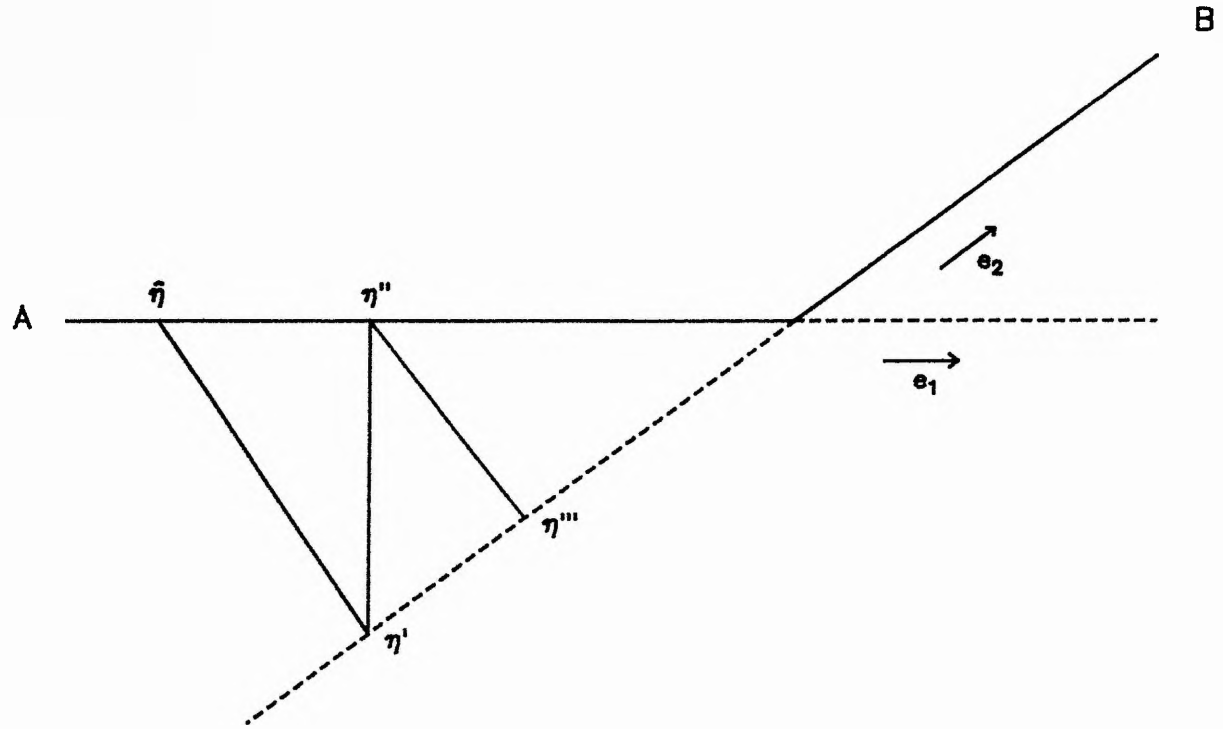


Fig. 7.11 The projections η', η'' and η'''.

since $\Lambda_d^N = 0$ if $\tau \in (-\pi,-\pi/2] \cup [\pi/2,\pi]$. A numerical integration routine is used to calculate Λ^N_{RMS}.

The maximum intrinsic pl-curvature Λ^N_{max} is obtained by maximising Λ_d^N over all angles τ and, from (7.17), is given by

$$\Lambda^N_{max} = \frac{\rho \sin(w^N)}{\xi(1 + \cos(w^N))} \tag{7.20}$$

This corresponds to $\tau = 0$, or equivalently, the half lifted line in direction e_1. It is intuitively sensible that the maximum will occur at this point (or at least, in the vicinity of this point) since values of τ close to zero correspond to short distances from $\hat{\eta}$ to C.

7.4.4.2 Parameter effects Pl-curvature

If $\tau \in (-\pi,-\pi/2] \cup [\pi/2,\pi]$, then $\Lambda_d^T = 0$. Otherwise, the calculation of the parameter effects pl-curvatures is simplified by first considering the half lifted line in direction e_1. This line hits B after travelling a distance ξ and can be written as

$$\eta^+_{e_1}(b) = \begin{cases} \hat{\eta} + be_1 & 0 \le b < \xi, \\ \hat{\eta} + \xi e_1 + (b-\xi)\kappa & \xi \le b, \end{cases}$$

where κ is a vector lying in rank (Y) (ie $\kappa = \lambda_1 e_B + \lambda_2 e_2$ for scalars λ_1 and λ_2). The projection of $\eta^+_{e_1}(b)$ onto A is given by

$$P_A\eta^+_{e_1}(b) = \begin{cases} \hat{\eta} + be_1 & 0 \le b < \xi, \\ \hat{\eta} + \xi e_1 + (b - \xi)k & \xi \le b, \end{cases} \tag{7.21a}$$

where

$$k = \lambda_1 \cos(w^N) e_1 + \lambda_2 e_2$$
$$= \alpha_1 e_1 + \alpha_2 e_2, \qquad (7.21b)$$

for scalars α_1 and α_2.

The half lifted line in direction d can be written in terms of κ as

$$\eta_d^+(b) = \begin{cases} \bar{\eta} + bd & 0 \le b < \xi_d, \\ \bar{\eta} + \xi_d d + (b - \xi_d)(\cos(\tau)\kappa + \sin(\tau)e_2) & \xi_d \le b \end{cases}$$

Thus, the projection of $\eta_d^+(b)$ onto A is

$$P_A \eta_d^+(b) = \begin{cases} \bar{\eta} + bd & 0 \le b < \xi_d, \\ \bar{\eta} + \xi_d d + (b - \xi_d) r_d & \xi_d \le b \end{cases} \qquad (7.22a)$$

where

$$r_d = \cos(\tau)k + \sin(\tau)e_2$$
$$= [k, e_2]g \qquad (7.22b)$$

The projection of $\eta_d^+(b)$ onto A consists of two line segments, one in direction d and the other in direction r_d. The angle w_d^G between the two line segments is given by

$$\cos(w_d^G) = d' r_d / |r_d|$$

$$= \frac{\alpha_1 \cos^2(\tau) + \alpha_2 \cos(\tau)\sin(\tau) + \sin^2(\tau)}{((\alpha_1^2 + \alpha_2^2)\cos^2(\tau) + 2\alpha_2 \cos(\tau)\sin(\tau) + \sin^2(\tau))^{\frac{1}{2}}}, \tag{7.23}$$

from which the geodesic pl-curvature is calculated using (7.14).

The speed of the projection of $\eta_d^+(b)$ onto $A \cap V_d$ is 1 before the line hits B and

$$v_d = |r_d| \cos(w_d^G)$$

$$= \alpha_1 \cos^2(\tau) + \alpha_2 \cos(\tau)\sin(\tau) + \sin^2(\tau) \tag{7.24}$$

afterwards, from which the parallel pl-curvature is calculated using (7.15). The relative parameter effects pl-curvature Λ_d^T is then calculated from (7.16).

To calculate the maximum and RMS parameter effects pl-curvatures Λ^T_{max} and Λ^T_{RMS}, it is sufficient to know α_1 and α_2, which are found as follows. The half lifted line in direction e_1 is the image on the solution locus of the line

$$\theta(b) = \bar{\theta} + bh^*$$

for some vector h^*. Hence

$$e_1 = Xh^*,$$

$$\kappa = Yh^*,$$

from which h^* and $\varkappa$ are found sequentially, since e_1, X and Y are known. Then k is calculated as

$$k = P_{\text{rank}(X)}\varkappa,$$

and α_1 and α_2 are found from the relationships

$$e_1'k = \alpha_1,$$

$$k'k = \alpha_1^2 + \alpha_2^2,$$

with the identifiability constraint that $\alpha_2 \geq 0$.

The RMS parameter effects pl-curvature Λ^T_{RMS} is given by

$$(\Lambda^T_{RMS})^2 = \frac{1}{S(2)} \int_{-\pi}^{\pi} (\Lambda^T_d)^2 \, d\tau,$$

where $S(2)$ is given by (7.19), and is calculated using a numerical integration routine.

The maximum parameter effects pl-curvature Λ^T_{max} is found by maximising $\Lambda_d{}^T$ over all angles τ. In general, a numerical maximisation routine must be used; a suitable initial value for such a routine is $\tau = 0$, since this corresponds to the shortest distance from $\hat{\eta}$ to C.

7.4.5 An Example

The maximum and RMS pl-curvatures are now found for model (7.2). Although the solution locus of this model is only 1-dimensional, the model serves to illustrate many of the principles discussed above.

The solution locus of model (7.2) consists of the sl-segments of two lines A and B, which join at the point C (Fig. 7.12). (Strictly, the point (1,1) constitutes a third (degenerate) line segment; however, it is sufficiently far away from the region of interest to be ignored.) To obtain the pl-curvatures in all directions, it is sufficient to consider the half lifted line in direction e_1 (towards C) and the half lifted line in direction $-e_1$ (away from C) (Fig. 7.12).

The half lifted line in direction $-e_1$ never hits another line segment so the relative intrinsic and parameter effects pl-curvatures in direction $-e_1$ are both zero.

The half lifted line in direction e_1 hits B at C after travelling a distance $\xi = \xi_{e_1} = 0.096$. The vectors e_1 and e_B are given by

$$e_1 = (0.316, 0.949)', \qquad e_B = (1, 0)',$$

so

$$\cos(w^N) = e_1' e_B = 0.316.$$

Hence, from (7.12) and (7.13) (or from (7.17)), the relative intrinsic pl-curvature in direction e_1 is $\Lambda_{e_1}{}^N = 0.751$.

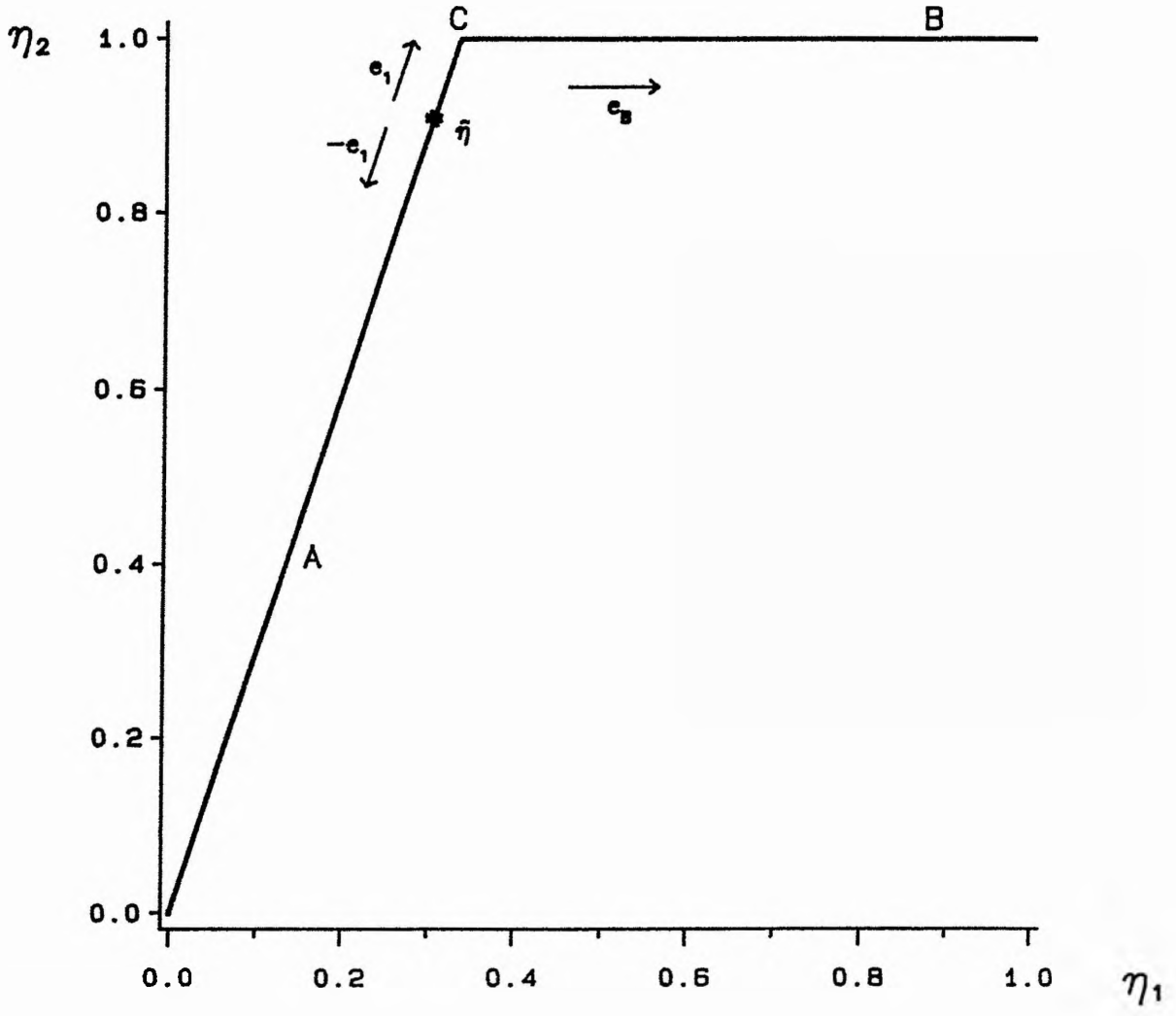


Fig. 7.12 The solution locus of model (7.2).

The maximum intrinsic pl-curvature is

$$\Lambda^N_{max} = \max\{0, 0.751\} = 0.751,$$

and, from (7.18), the RMS intrinsic pl-curvature is

$$\Lambda^N_{RMS} = \sqrt{\{(0^2 + 0.751^2)/2\}},$$
$$= 0.531$$

Since $1/(2\sqrt{F(1,\infty;0.05)}) = 0.255$, the intrinsic pl-curvatures indicate that the tangent plane approximation to the solution locus is very poor in the region of interest.

The relative parameter effects pl-curvature in direction e_1 is obtained as follows. The projection of the half lifted line in direction e_1 onto the tangent line at $\hat{\eta}$ (ie the line A) is identical to the tangent line. Since this projection does not change direction at C, the geodesic pl-curvature is zero. In this 1-dimensional example, it is easier to calculate the parallel pl-curvature by returning to first principles rather than by using the techniques described in Section 7.4.4.2. The sl-segments of A and B are given by

$$sl(A) = \{\eta(\theta) = (\theta, 3\theta); \theta \leq 1/3\}$$

$$sl(B) = \{\eta(\theta) = (\theta, 1); 1/3 < \theta \leq 1\}.$$

Hence, the speed of a particle travelling along the half lifted line is $\sqrt{10}$ on sl(A) and 1 on sl(B). The speed of the projection of the half lifted line onto the tangent line is then

$$v = 1\cos(w^N) = 0.316$$

and from (7.15), the parallel pl-curvature in direction e_1 is

$$c^P_{e_1} = \frac{0.316 - 3.16}{0.096(3.16 + \sqrt{(0.316^2 - 2*0.316*3.16 + 2*3.16^2)})}$$

$$= 4.00.$$

Thus, from (7.16), the relative pl-curvature in direction e_1 is

$$\Lambda^T_{e_1} = \rho\, c^P_{e_1} = 0.40.$$

Hence,

$$\Lambda^T_{max} = 0.40,$$

$$\Lambda^T_{RMS} = 0.28,$$

which are large compared to $1/(2\sqrt{F})$, so the parameter effects pl-curvatures indicate that the uniform co-ordinates approximation is very poor.

The pl-curvatures are large and indicate that any inference based on a linear approximation is likely to give a very misleading result. This is demonstrated by the comparison of the exact and the linear approximation 95% confidence intervals for Θ given in Section 7.1.

7.5 CASE 2: MULTI-HYPERPLANAR SOLUTION LOCUS

The more general case of a model with a multi-hyperplanar solution locus is now considered. Although the dimension of the solution locus has increased, the study of the behaviour of the half lifted lines still reduces to a problem in three dimensions.

Suppose that the solution locus in the region of interest about $\hat{\eta}$ consists of the sl-segments of r+1 p-dimensional hyperplanes A, B_1,...,B_r. Suppose that $\hat{\eta}$ lies on the interior of sl(A) and that sl(A) meets sl(B_i) in a (p-1)-dimensional hyperplane C_i, i = 1...r. The sl-segment of A can be regarded as being bordered by the sl-segments of B_i, i = 1...r.

The half lifted line $\eta_h^+(b)$ either

i) travels outwards from $\hat{\eta}$ at constant speed in a straight line on sl(A) to infinity,

ii) travels at constant speed in a straight line on sl(A) until it hits one of the hyperplanes B_1,...,B_r, say B_h, and then travels in a straight line on sl(B_h) in a new direction and at a new constant speed. (In general, $\eta_h^+(b)$ hits a large number of hyperplanes, of which B_h is only the first. However, it is assumed that only the first two segments of $\eta_h^+(b)$ (ie the ones on sl(A)

and sl(B_h)) are in the region of interest and that the others can be effectively ignored. This assumption is discussed further in Section 7.8.)

The nonlinearity of η_h^+ (b) can be split into three components, as before. Let U_h be the 2-dimensional plane that contains the projection of η_h^+ (b) onto A; thus, in the simple case described in Section 7.4, U_h is the tangent plane A. If V_h is the 2-dimensional plane orthogonal to A and containing $\dot{\eta}_h$, the three components are:

i) the normal component which measures the effect of the change in direction normal to A (ie in V_h),

ii) the geodesic component which measures the effect of the change of direction in A (ie in U_h),

iii) the parallel component which measures the effect of the change of speed of the particle in the line $U_h \cap V_h$.

The pl-curvatures are essentially the same as before; only the definitions of the angles w_h^N, w_h^G and the speed v_h are slightly altered. For example, the projection of η_h^+ (b) onto V_h consists of two line segments, one lying in sl(A), the other in sl(B_h), with an angle $w_h^N \in (0,\pi/2]$ between them; the intrinsic pl-curvature is then given by (7.12). Similarly, the projection of η_h^+ (b) onto U_h consists of two line segments with an angle $w_h^G \in (0,\pi/2]$ between them; the geodesic pl-curvature is then given by (7.14). Finally, the parallel pl-curvature is given by (7.15) where $\|\dot{\eta}_h\|$ and v_h are the two speeds of the projection of η_h^+ (b) onto $U_h \cap V_h$.

The curvature in every possible direction is obtained by considering the half lifted lines that head in direction d, for all unit vectors d in rank(X).

7.5.1 Calculating PI-Curvatures

Let $Y_1,...,Y_r$ be the matrices associated with $B_1,...,B_r$ respectively. Since A and B_i intersect in a $(p-1)$-dimensional hyperplane, there exist vectors e_{i1}, e_{iB}, e_{i2}, $e_{i3},...,e_{ip}$, such that e_{i1}, e_{i2}, $e_{i3},...,e_{ip}$ is an orthonormal basis of rank(X), e_{iB}, e_{i2}, $e_{i3},...,e_{ip}$ is an orthonormal basis of rank(Y_i), $i = 1...r$.

Let E_i be the nxp matrix

$$E_i = [e_{i1}, e_{i2}, ..., e_{ip}], \quad i = 1...r,$$

and let L_i be a pxp orthogonal matrix such that

$$E_i = E_1 L_i, \quad i = 2...r.$$

Let $w_i^N \in (0,\pi/2]$ be the angle between A and B_i and let ξ_i be the shortest distance from $\hat{\eta}$ to sl(B_i).

Given E_1, the unit vectors d in rank(X) are parameterised by a $(p-1)$-vector of angles τ as

$$d = d(\tau) = E_1 g(\tau).$$

where

$$g(\tau) = \begin{pmatrix} g_1(\tau) \\ g_2(\tau) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ g_{p-1}(\tau) \\ g_p(\tau) \end{pmatrix} = \begin{pmatrix} \cos(\tau_1) \\ \sin(\tau_1)\cos(\tau_2) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \sin(\tau_1)\sin(\tau_2)\ldots\cos(\tau_{p-1}) \\ \sin(\tau_1)\sin(\tau_2)\ldots\sin(\tau_{p-1}) \end{pmatrix}$$

and where $\tau_1,\ldots,\tau_{p-2} \in [0,\pi)$ and $\tau_{p-1} \in (-\pi,\pi]$. If a half lifted line hits B_i, i = 2...r, then d can be written in terms of the orthonormal basis E_i as

$$d = E_1 g(\tau) = E_i L_i' g(\tau) = E_i g(\psi_i)$$

for a vector of angles $\psi_i = (\psi_{i1},\ldots,\psi_{i,p-1})'$, i = 2...r.

7.5.2 Intrinsic Pl-Curvature

Suppose that η_d^+ (b) hits B_1. The projection of η_d^+ (b) onto V_d consists of two line segments, one in direction d, the other in direction d^N, where d^N is a unit vector given by

$$d^N = [e_{1B}, e_{12},\ldots,e_{1p}]g^N(\tau)$$

for a p-vector

$$g^N(\tau) = (g_1^N(\tau),\ldots,g_p^N(\tau))'.$$

The projection of d^N onto rank(X) is parallel to d (cf Section 7.4.4.1) so that

$$g^N(\tau) = \frac{1}{\delta}\left(\frac{g_1(\tau)}{\cos(w_1^N)}, g_2(\tau), g_3(\tau), \ldots, g_p(\tau)\right)'$$

where

$$\delta^2 = \frac{1-\sin^2(\tau_1)\sin^2(w_1^N)}{\cos^2(w_1^N)}.$$

Hence

$$\cos(w_d^N) = d'd^N = 1/\delta,$$

$$\xi_d = \xi_1/\cos(\tau_1),$$

so that, from (7.12) and (7.13),

$$\Lambda_d^N = \frac{\rho\cos^2(\tau_1)\sin(w_1^N)}{\xi_1(\cos(w_1^N)+\sqrt{\{1-\sin^2(\tau_1)\sin^2(w_1^N)\}})}$$

If $\eta|_{d^+}$ (b) hits B_i, for $i \neq 1$, then

$$d = E_1 g(\tau) = E_i g(\psi_i)$$

so that

$$d^N = [e_{iB}, e_{i2}, \ldots, e_{ip}]g^N(\psi_i)$$

etc. Thus

$$\Lambda_d^N = \frac{\rho \cos^2(\psi_{i1})\sin(w_i^N)}{\xi_i(\cos(w_i^N) + \sqrt{(1 - \sin^2(\psi_{i1})\sin^2(w_i^N))})}$$

which can be written in terms of τ by using the relationship

$$\cos(\psi_{i1}) = [L_i' g(\tau)]_1.$$

To calculate Λ^N_{max} and Λ^N_{RMS}, it is sufficient to know e_{i1}, e_{iB}, $i = 1...r$, and to relate e_{i1}, $i = 2...r$, to E_1. As before, e_{i1} and e_{iB} are given by

$$e_{i1} = (\eta_i'' - \hat{\eta})/|\eta_i'' - \hat{\eta}|.$$

$$e_{iB} = (\eta_i''' - \eta_i')/|\eta_i''' - \eta_i'|$$

where

$$\eta_i' = P_{B_i}\hat{\eta} \quad , \quad \eta_i'' = P_A\eta_i' \quad , \quad \eta_i''' = P_{B_i}\eta_i''.$$

Then

$$\cos(w_i^N) = e_{i1}' e_{i\beta},$$

$$\xi_i = |\eta_i' - \hat{\eta}|/\sin(w_i^N).$$

The vectors e_{i1}, $i = 2...r$, are sequentially related to E_1 as follows. The vectors $e_{12},e_{13},...,e_{1p}$ are arbitrary orthogonal unit vectors in rank(X) $\cap$ rank(Y). Thus, without loss of generality

$$e_{21} = l_{21} e_{11} + l_{22} e_{12}$$

for scalars l_{21}, l_{22} given by

$$e_{21}' e_{11} = l_{21},$$

$$e_{21}' e_{21} = l_{21}^2 + l_{22}^2 = 1,$$

with the identifiability constraint that $l_{22} \geq 0$. Hence,

$$e_{12} = (e_{21} - l_{21} e_{11})/l_{22}.$$

Thus, L_2 is taken to be of the form

$$L_2 = \begin{pmatrix} l_{21} & \\ l_{22} & \\ 0 & \\ . & L_2^* \\ . & \\ . & \\ 0 & \end{pmatrix}$$

where the px(p-1) matrix L_2^* remains undefined. (If $p = 1$, then $e_{21} = -e_{11}$. If $p > 1$ and $l_{22} = 0$, then it is not necessary to find e_{12} explicitly. In the subsequent development, it is assumed that the problem is not degenerate in this way).

This process is repeated, until all the e_{i1} are related to E_1. At each stage, another vector of the orthonormal basis E_1 is defined. For example, without loss of generality

$$e_{31} = l_{31}e_{11} + l_{32}e_{12} + l_{33}e_{13}$$

for scalars l_{31}, l_{32} and l_{33}. In general, at the ith stage

$$e_{i1} = \sum_{j=1}^{i} l_{ij}e_{1j},$$

so that L_i is of the form

$$L_i = \begin{pmatrix} l_{i1} & \\ l_{i2} & \\ \cdot & \\ \cdot & \\ \cdot & \\ l_{ii} & L_i^* \\ 0 & \\ \cdot & \\ \cdot & \\ \cdot & \\ 0 & \end{pmatrix}$$

The scalars l_{ij}, $j = 1...i$, and the vector e_{1i} are found from the relationships

$$e_{i1}'e_{1j} = l_{ij}, \qquad j = 1...(i-1),$$

$$\sum_{j=1}^{i} l_{ij}^2 = 1,$$

with the constraint $l_{ii} > 0$, and

$$e_{ii} = (e_{ii} - \sum_{j=1}^{i-1} l_{ij} e_{ij})/l_{ii}.$$

In general, r vectors of the orthonormal basis E_1 are found.

The RMS intrinsic pl-curvature is given by

$$(\Lambda^N_{RMS})^2 = \frac{1}{S(p)} \int_0^{\pi} d\tau_1 (\sin(\tau_1))^{p-2} \int_0^{\pi} d\tau_2 (\sin(\tau_2))^{p-3} \ldots \int_0^{\pi} d\tau_{p-2} \sin(\tau_{p-2}) \int_{-\pi}^{\pi} d\tau_{p-1} (\Lambda^N_d)^2.$$

where $S(p)$ is given by (7.19). If $\eta_d{}^+ (b)$ hits $B_i, \Lambda_d{}^N$ is a function of the angle ψ_{i1} and hence of the angles $\tau_1,\ldots, \tau_i$ by the construction of the matrix L_i. Thus, the intrinsic pl-curvature of any half lifted line only depends on the angles $\tau_1,\ldots,\tau_r$ at most; hence, if $r < (p - 1)$, the integral above can be simplified to

$$(\Lambda^N_{RMS})^2 = \frac{S(p-r)}{S(p)} \int_0^{\pi} d\tau_1 (\sin(\tau_1))^{p-2} \ldots \int_0^{\pi} d\tau_r (\sin(\tau_r))^{p-r-1} (\Lambda^N_d)^2.$$

This is calculated using a numerical integration routine.

The maximum intrinsic pl-curvature can usually be found as follows. Suppose that the hyperplanes $B_2,\ldots,B_r$ are "removed" so that the solution locus consists of just sl(A) and sl(B_1). The maximum intrinsic pl-curvature for this solution locus is called the maximum intrinsic pl-curvature associated with B_1 and is given by

$$(\Lambda^N_{max})_1 = \frac{\rho \sin(w_1^N)}{\xi_1(1 + \cos(w_1^N))} \qquad (7.25)$$

from (7.20); this corresponds to the half lifted line in direction e_1. When the other hyperplanes are "replaced", the maximum intrinsic pl-curvature of all the half lifted lines that hit B_1 is bounded above by $(\Lambda^N_{max})_1$, with equality if the half lifted line in direction e_1 hits B_1 (rather than another hyperplane). If the bound is attained, $(\Lambda^N_{max})_1$ is called an exact maximum. This process is repeated for all the hyperplanes B_i, $i = 1...r$, in turn. Then,

$$\Lambda^N_{max} \leq \max_{i=1...r} (\Lambda^N_{max})_i = (\Lambda^N_{max})^*$$

with equality if $(\Lambda^N_{max})^*$ is exact. Usually, $\Lambda^N_{max} = (\Lambda^N_{max})^*$ since in most cases, $(\Lambda^N_{max})^*$ corresponds to one of the sl-segments closest to $\hat{\eta}$, and these sl-segments often give exact maxima.

If the bound is not attained, then $(\Lambda^N_{max})^*$ is still a very useful measure of intrinsic nonlinearity. If $(\Lambda^N_{max})^*$ is small, Λ^N_{max} is small and the tangent plane approximation will be good. If $(\Lambda^N_{max})^*$ is large, it is likely that Λ^N_{max} is large and inference using the tangent plane approximation should only be used with caution. If Λ^N_{max} is required exactly, a numerical maximisation procedure must be used.

7.5.3 Parameter Effects Pl-Curvature

Without loss of generality, the half lifted line in direction e_{i1} can be written

$$\eta^{+}_{e_{i1}}(b) = \begin{array}{ll} \hat{\eta} + be_{i1} & 0 \le b < \xi_i \\ \hat{\eta} + \xi_i e_{i1} + (b - \xi_i)\kappa_i & \xi_i \le b \end{array}$$

where

$$\kappa_i = \lambda_{i1} e_{i8} + \lambda_{i2} e_{i2}.$$

Thus,

$$P_{\hat{\eta}}\, \eta^{+}_{e_{i1}}(b) = \begin{array}{ll} \hat{\eta} + be_{i1} & 0 \le b < \xi_i \\ \hat{\eta} + \xi_i e_{i1} + (b - \xi_i)k_i & \xi_i \le b \end{array}$$

where

$$k_i = \lambda_{i1} \cos(\omega_i^N) e_{i1} + \lambda_{i2} e_{i2} = \alpha_{i1} e_{i1} + \alpha_{i2} e_{i2},$$

(cf equations (7.21)).

Suppose $\eta_d^+(b)$ hits B_i. Then

$$\eta^{+}_{d}(b) = \begin{array}{ll} \hat{\eta} + bE_i g(\psi_i) & 0 \le b < \xi_d \\ \hat{\eta} + \xi_d E_i g(\psi_i) + (b - \xi_d)[\kappa_i, e_{i2}, e_{i3}, \ldots, e_{ip}] g(\psi_i) & \xi_d \le b \end{array}$$

(where $\Psi_i = \tau$ if $i = 1$). Thus

$$P_{\Lambda}\eta_d^*(b) = \begin{cases} \hat{\eta} + bE_1 g(\psi_1) & 0 \le b < \xi_d \\ \hat{\eta} + \xi_d E_1 g(\psi_1) + (b - \xi_d) r_d & \xi_d \le b \end{cases}$$

where

$$r_d = [k_1, e_{12}, \ldots, e_{1p}] g(\psi_1)$$

$$= \alpha_{11} g_1(\psi_1) e_{11} + (\alpha_{12} g_1(\psi_1) + g_2(\psi_1)) e_{12} + \sum_{j=3}^{r} g_j(\psi_1) e_{1j},$$

(cf equations (7.22)). Then

$$v_d = r_d' d$$

$$= \alpha_{11} \cos^2(\psi_{11}) + \alpha_{12} \cos(\psi_{11}) \sin(\psi_{11}) \cos(\psi_{12}) + \sin^2(\psi_{11}),$$

(cf equation (7.24),) from which the parallel pl-curvature is calculated using (7.15). Also,

$$|r_d|^2 = (\alpha_{11}^2 + \alpha_{12}^2) \cos^2(\psi_{11}) + 2\alpha_{12} \cos(\psi_{11}) \sin(\psi_{11}) \cos(\psi_{12}) + \sin^2(\psi_{11})$$

and

$$\cos(w_d^G) = v_d / |r_d|$$

(cf equation (7.23),) from which the geodesic pl-curvature is calculated using (7.14). The relative parameter effects pl-curvature is then calculated from (7.16).

Since $\Lambda_d{}^T$ depends on both ψ_{i1} and ψ_{i2}, $\Lambda^T{}_{max}$ and $\Lambda^T{}_{RMS}$ can be calculated if α_{i1}, α_{i2}, e_{i1} and e_{i2}, $i = 1...r$, are known and if e_{i1} and e_{i2} are sequentially related to E_1, $i = 2...r$. The values of α_{i1} and α_{i2} are found as before (Section 7.4.4.2) and the vectors e_{i2} are then given by

$$e_{i2} = (k_i - \alpha_{i1} e_{i1})/\alpha_{i2}, \qquad i = 1...r.$$

The vectors e_{i1} and e_{i2} are sequentially related to E_1 by taking the matrices L_i to be of the form

$$L_i = \begin{pmatrix} l_{i11} & l_{i12} \\ l_{i21} & l_{i22} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ l_{i,2i-1,1} & l_{i,2i-1,2} \\ 0 & l_{i,2i,2} \\ 0 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 0 \end{pmatrix} L_i^*$$

At each stage, the scalars l_{ijk} and the vectors $e_{1,2i-1}$ and $e_{1,2i}$ are found from the relationships

$$L_i = E_1' E_i,$$

$$L_i' L_i = I_p,$$

with the identifiability constraints that $l_{i,2i-1,1} > 0$ and $l_{i,2i,2} > 0$. In general, $2r$ vectors of the orthonormal basis E_1 are explicitly constructed. It is important to note that in general, the matrices $\{L_i\}$ and the vectors of E_1 used to calculate the parameter effects pl-curvatures are different from those used to calculate the intrinsic pl-curvature.

The RMS parameter effects pl-curvature is given by

$$(\Lambda^T_{RMS})^2 = \frac{1}{S(p)} \int_o^{\pi} d\tau_1 (\sin(\tau_1))^{p-2} \ \ldots \ \int_o^{\pi} d\tau_{p-2} \sin(\tau_{p-2}) \int_{-\pi}^{\pi} d\tau_{p-1} (\Lambda_d^T)^2$$

and if $2r < (p - 1)$, this simplifies to

$$(\Lambda^T_{RMS})^2 = \frac{S(p-2r)}{S(p)} \int_o^{\pi} d\tau_1 (\sin(\tau_1))^{p-2} \ldots \int_o^{\pi} d\tau_{2r} (\sin(\tau_{2r}))^{p-2r-1} (\Lambda_d^T)^2$$

The maximum parameter effects pl-curvature is found by calculating the maximum parameter effects pl-curvature $(\Lambda^T_{max})_i$ associated with each hyperplane B_i in turn, $i = 1 \ldots r$. Suppose all the hyperplanes except B_i are removed and let $(\Lambda_d{}^T)_i$ be the parameter effects pl-curvature of this "reduced" solution locus. If $p > 2$, $(\Lambda^T_{max})_i$ is found by maximising $(\Lambda_d{}^T)_i$ over two angles ψ_{i1} and ψ_{i2}. A suitable initial value for ψ_{i1} is $\psi_{i1} = 0$, since this corresponds to the shortest distance from $\hat{\eta}$ to $sl(B_i)$. Also, $(\Lambda_d{}^T)_i|_{\psi_{i1}=0}$ is independent of ψ_{i2} and

$$\frac{d(\Lambda_d^T)_i}{d\psi_{i1}} \Big|_{\psi_{i1}=0} = \Delta \cos(\psi_{i2})$$

where Δ is a constant dependent on α_{i1}, α_{i2} and ξ_i. Thus, a suitable initial value for ψ_{i2} is $\psi_{i2} = 0$ if $\Delta > 0$ and $\psi_{i2} = \pi$ if $\Delta < 0$, since incrementing values of (ψ_{i1}, ψ_{i2}) along the line $\psi_{i2} = 0$ if $\Delta > 0$ or $\psi_{i2} = \pi$ if $\Delta < 0$ will give a steep initial increase in values of $(\Lambda_d{}^T)_i$; the sign of Δ is easily calculated by a numerical derivative. A typical example of the form of $(\Lambda_d{}^T)_i$ as a function of ψ_{i1} and ψ_{i2} is shown in Figure 7.13. In this case, $\Delta > 0$, and the global maximum occurs along the line $\psi_{i2} = 0$. In fact,

$$\frac{d(\Lambda_d^T)_i}{d\psi_{i2}} \Big|_{\psi_{i2}=0,\pi} = 0$$

and in all the examples considered, the global maximum has occured along the line $\psi_{i2} = 0$ if $\Delta > 0$ and $\psi_{i2} = \pi$ if $\Delta < 0$, although it has not been possible to prove that this is true in general. Let d_i be the direction of the half lifted line such that

$$(\Lambda_{d_i}^T)_i = (\Lambda_{max}^T)_i.$$

Then, $(\Lambda^T{}_{max})$ is exact if $\eta_{d_i}{}^+(b)$ still hits B_i when the other hyperplanes are replaced. Thus, $\Lambda^T{}_{max}$ is bounded above by $(\Lambda^T{}_{max})^*$ where

$$(\Lambda_{max}^T)^* = \max_{i=1\ldots r} (\Lambda_{max}^T)_i$$

with equality if $(\Lambda^T{}_{max})^*$ is exact. As with the maximum intrinsic pl-curvature, this bound has always been attained in theexamples considered.

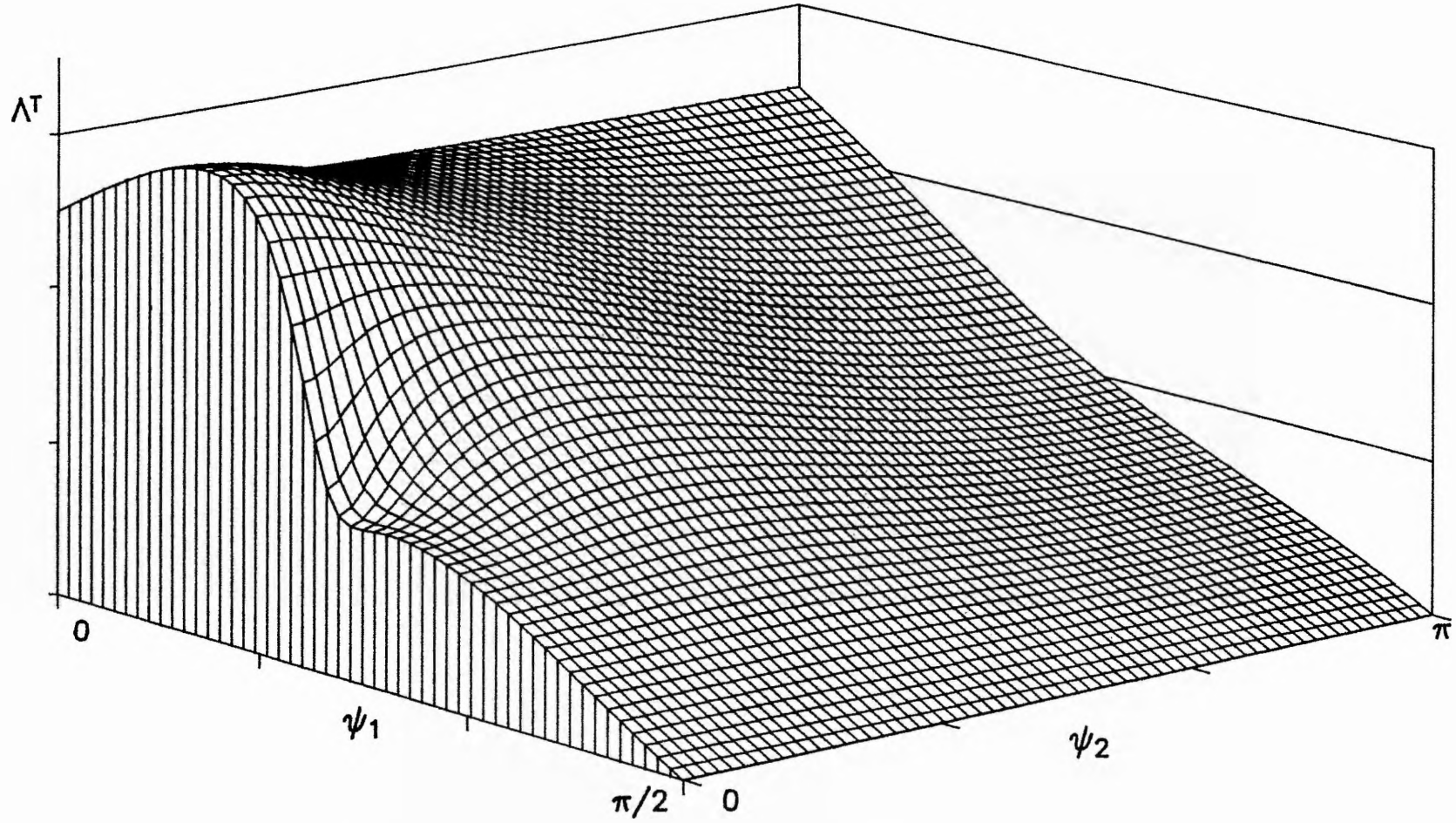


Fig. 7.13 The typical form of Λ^T as a function of ψ_1 and ψ_2.

7.6 AN EXAMPLE: THE GROWTH RATE OF BROWN TROUT FED ON MAXIMUM RATIONS

The growth rate $\bar{g}$ of brown migratory trout weighing approximately 50 g and fed on maximum rations depends on the water temperature T (Elliott, 1975). This relationship is well described by the nonlinear change-point model

$$\bar{g}_i = \begin{cases} \psi_M + b_1(T_i - T_M) + \epsilon_i & \text{if } T_i \leq T_M \\ \psi_M + b_2(T_i - T_M) + \epsilon_i & \text{if } T_i > T_M, \end{cases} \tag{7.26}$$

where $\{\epsilon_i\}$ are NID$(0,\sigma^2)$ (Section 8.3). The growth model has four parameters, ψ_M, T_M, b_1, b_2, where, in particular, T_M is the temperature at which maximum growth occurs and ψ_M is the maximum growth rate. The parameter T_M defines an (unknown) change point.

Model (7.26) can be reparameterised as

$$\bar{g}_i = \begin{cases} \theta_1 + \theta_2 T_i + \epsilon_i & \text{if } T_i \leq (\theta_3 - \theta_1)/(\theta_2 - \theta_4) \\ \theta_3 + \theta_4 T_i + \epsilon_i & \text{if } T_i > (\theta_3 - \theta_1)/(\theta_2 - \theta_4), \end{cases} \tag{7.27}$$

where

$$\begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix} = \begin{pmatrix} \psi_M - b_1 T_M \\ b_1 \\ \psi_M - b_2 T_m \\ b_2 \end{pmatrix}.$$

Parameterised by θ, the growth model is continuous and piecewise linear, since

i) $\theta_1 + \theta_2 T$ and $\theta_3 + \theta_4 T$ are linear in θ,

ii) the model function

$$f(T,\theta) = (\theta_1 + \theta_2 T)\mathrm{I}\{T \le (\theta_3 - \theta_1)/(\theta_2 - \theta_4)\} + (\theta_3 + \theta_4 T)\mathrm{I}\{T > (\theta_3 - \theta_1)/(\theta_2 - \theta_4)\}$$

is continuous in Θ. (It is assumed that there is a genuine change point, so that $b_1 \neq b_2$ and thus $\theta_2 \neq \theta_4$).

Hence, the solution locus is continuous and piecewise hyperplanar and on each sl-segment, the Θ parameter lines are straight, parallel and equispaced. Thus, the pl-curvatures can be used to measure the intrinsic nonlinearity of the model and the parameter effects nonlinearity of the parameterisation (7.27).

Values of T_i and $\bar{g}_i$, i = 1...15, obtained from a series of experiments (Elliott, 1975, Section 8.1) are given in Table 7.1 and are ordered so that

$$T_1 < T_2 < \ldots < T_{15}.$$

Each sl-segment of the solution locus can then be written as

TABLE 7.1

Temperature and growth rate data

T	$\bar{g}$
3.8	0.020
4.2	0.062
5.6	0.169
6.8	0.254
7.1	0.351
9.5	0.499
10.8	0.592
12.8	0.747
13.6	0.743
15.0	0.616
16.2	0.418
17.8	0.216
19.5	0.010
20.4	-0.217
21.7	-0.367

$$\left\{ \eta(\theta) = \begin{pmatrix} 1 & T_1 & & \\ 1 & T_2 & & \\ \cdot & \cdot & & \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & & \\ 1 & T_j & & \\ & & 1 & T_{j+1} \\ & & \cdot & \\ 0 & 0 & \cdot & \\ & & \cdot & \\ & & 1 & T_{15} \end{pmatrix} \theta ; \quad \theta \in \Theta_j \right\}$$

where

$$\Theta_j = \{\theta ; T_j \leq (\theta_3 - \theta_1)/(\theta_2 - \theta_4) < T_{j+1}\}$$

$$= \{\theta ; T_j \leq T_M(\theta) < T_{j+1}\}$$

for some integer $j \in \{1...14\}$.

The least squares estimate of T_M is $\hat{T}_M = 13.4$, so $\hat{\eta}$ lies on the sl-segment of hyperplane A where

$$A = \left\{ \begin{pmatrix} 1 & 3.8 & & \\ \cdot & \cdot & & \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & & \\ 1 & 12.8 & & \\ & & 1 & 13.6 \\ & & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot \\ & & \cdot & \cdot \\ & & 1 & 21.7 \end{pmatrix} \theta ; \quad \theta \in R^4 \right\}.$$

The solution locus "changes" hyperplane when T_M equals one of the temperature design points $\{T_i\}$, so sl(A) is bordered by the sl-segments of two hyperplanes B_1 and B_2 where

$$B_1 = \left\{ \begin{pmatrix} 1 & 3.8 & & \\ \vdots & \vdots & & \\ 1 & 10.8 & 0 & 0 \\ 0 & 0 & 1 & 12.8 \\ & & \vdots & \vdots \\ & & 1 & 21.7 \end{pmatrix} \theta ; \quad \theta \in R^4 \right\},$$

$$B_2 = \left\{ \begin{pmatrix} 1 & 3.8 & & \\ \vdots & \vdots & & \\ 1 & 13.6 & 0 & 0 \\ 0 & 0 & 1 & 15.0 \\ & & \vdots & \vdots \\ & & 1 & 21.7 \end{pmatrix} \theta ; \quad \theta \in R^4 \right\}.$$

The hyperplanes intersect in 3-dimensional hyperplanes C_1 and C_2 where

$$sl(C_1) = \{\eta(\theta) ; T_M = 12.8\} = \{\eta\{\theta) ; 12.8(\theta_2 - \theta_4) = \theta_3 - \theta_1\}$$

$$sl(C_2) = \{\eta(\theta) ; T_M = 13.6\} = \{\eta\{\theta) ; 13.6(\theta_2 - \theta_4) = \theta_3 - \theta_1\}$$

A linear approximation 95% confidence interval for T_M is

$$T_M \in (13.0,\ 13.8)$$

which suggests that, even if the linear approximation is not very good, the 95% region of interest about $\hat{\eta}$ is contained in the sl-segments of A, B_1 and B_2.

The distances from $\hat{\eta}$ to sl(B_1) and sl(B_2) are

$$\xi_1 = 0.130,$$

$$\xi_2 = 0.038.$$

Since ξ_1 is greater than $\rho\sqrt{F(11,4;0.005)} = 0.120$, it is unlikely that the change of hyperplane from A to B_1 will cause large nonlinear behaviour. However, ξ_2 is small compared to $\rho\sqrt{F}$, so the change of hyperplane from A to B_2 could cause considerable nonlinearity.

7.6.1 Intrinsic Pl-Curvature

The angles between A and B_1 and between A and B_2 are

$$w_1^N = 56.3°,$$

$$w_2^N = 55.2°.$$

Thus, the solution locus deviates considerably at C_1 and C_2. The maximum intrinsic pl-curvature is found by calculating the maximum intrinsic pl-curvatures associated with B_1 and B_2 from equation (7.25)

$$(\Lambda_{max}^N)_1 = 0.27,$$

$$(\Lambda_{max}^N)_2 = 0.91.$$

Since the half lifted line in direction e_{21} hits B_2 (rather than B_1), $(\Lambda^N_{max})_2$ is exact and

$$\Lambda^N_{max} = 0.91.$$

Also, the RMS intrinsic pl-curvature is calculated to be

$$\Lambda^N_{RMS} = 0.26.$$

The maximum intrinsic pl-curvature is large compared to $1/(2\sqrt{F}) = 0.27$. This is as expected, because the angle between A and B_2 is large and because the distance from $\hat{\eta}$ to sl(B_2) is small. The RMS intrinsic pl-curvature is much smaller than the maximum pl-curvature (this is quite common in models with a comparatively large number of parameters) but is still close to the "critical value" of $1/(2\sqrt{F})$. Hence, the tangent plane approximation to the solution locus is poor.

7.6.2 Parameter Effects Pl-Curvature

The scalars α_{i1} and α_{i2} corresponding to B_1 and B_2 are

$$\alpha_{11} = 0.49, \quad \alpha_{12} = 0.55,$$

$$\alpha_{21} = 0.53, \quad \alpha_{22} = 0.54.$$

Thus,

$$v_1 = \alpha_{11} = 0.49,$$

$$w_1^c = \cos^{-1}(\alpha_{11}/\sqrt{(\alpha_{11}^2 + \alpha_{12}^2)}) = 48.3°,$$

and

$$v_2 = 0.53,$$

$$w_2^c = 45.5°,$$

so that the speeds and directions of the parameter lines in directions e_{11} and e_{21} change considerably at C_1 and C_2.

Maximising the parameter effects pl-curvature associated with B_1 and B_2 over the angles ψ_{i1} and ψ_{i2} gives

$$(\Lambda^T_{max})_1 = 0.26 \text{ corresponding to } \psi_{11} = 5.2°, \psi_{12} = 180.0°,$$

$$(\Lambda^T_{max})_2 = 0.84 \text{ corresponding to } \psi_{21} = 5.2°, \psi_{22} = 180.0°.$$

Since $(\Lambda^T_{max})_2$ is exact, $\Lambda^T_{max} = 0.84$. Also, $\Lambda^T_{RMS} = 0.19$. The maximum parameter effects pl-curvature is much larger than $1/(2\sqrt{F})$ so the uniform co-ordinates approximation is poor.

7.6.3 Large Nonlinear Behaviour

The pl-curvatures calculated above indicate that the nonlinear behaviour of model (7.27) is large. This can be demonstrated in a number of ways. For example, a 95% confidence interval for θ_3 obtained from a linear approximation is smaller than that obtained from a simulation (Table 7.2). Also, the simulated distributions of $\hat{\theta}_3$ and $\hat{\theta}_4$ are skewed (Table 7.2). However, perhaps the best way is by studying the simulated distribution of $(\hat{\theta}_3 - \hat{\theta}_1 - 13.6(\hat{\theta}_2 - \hat{\theta}_4))$, (Fig. 7.14); this is clearly non-normal and shows that the joint distribution of $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4)$ is far-from-linear. (The value 13.6 is used because it is the temperature design point closest to the least squares estimate $\hat{T}_M = 13.4$).

It is not clear how the large nonlinear behaviour of model (7.27) affects inferences in general. Certainly, results obtained from a linear approximation should only be used with caution. Probably the simplest altervative is to make inferences based on simulations (cf Section 3.4).

7.6.4 The Nonlinear Behaviour of a Reduced Data Set

The large nonlinearity of model (7.27) is mainly due to the half lifted lines that hit B_2. It is interesting to observe the effect of omitting the data point corresponding to $T = 13.6$. The solution locus in the region of interest now consists of the sl-segments of hyperplanes A, B_1 and B_2, where, for example

TABLE 7.2

Simulation results for model (7.27) - complete data set

Parameter	Lin approx 95% conf. int.		Simulation 95% conf. int.	
θ_1	-0.33	-0.21	-0.33	-0.21
θ_2	0.073	0.088	0.073	0.089
θ_3	2.53	2.85	2.53	2.90
θ_4	-0.149	-0.131	-0.151	-0.131

Parameter	θ_1	θ_2	θ_3	θ_4
Skewness stat	0.34	0.62	3.26**	-2.69**
Excess kurtosis stat	-0.91	-0.70	1.87	1.45

** indicates serious nonlinear behaviour

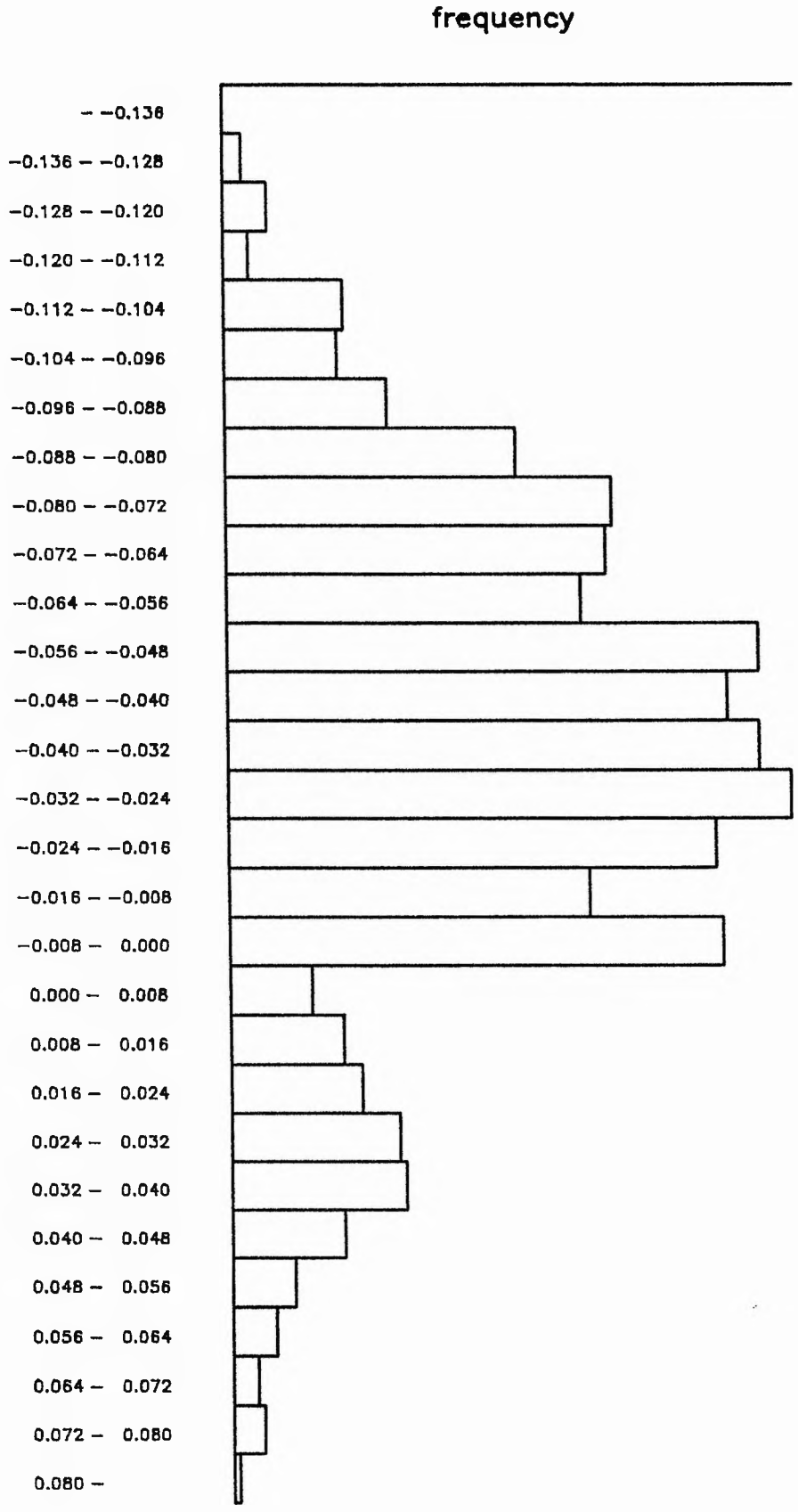


Fig. 7.14 The simulated distribution of $\hat{\theta}_3 - \hat{\theta}_1 - 13.6(\hat{\theta}_2 - \hat{\theta}_4)$.

$$A = \left\{ \begin{pmatrix} 1 & 3.8 & & \\ \vdots & \vdots & & \\ \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & & \\ 1 & 12.8 & & \\ & & 1 & 15.0 \\ & & \vdots & \vdots \\ 0 & 0 & \vdots & \vdots \\ & & \vdots & \vdots \\ & & 1 & 21.7 \end{pmatrix} \theta; \quad \theta \in R^4 \right\}.$$

The distances from $\hat{\eta}$ to sl(B_1) and sl(B_2) are

$$\xi_1 = 0.139,$$

$$\xi_2 = 0.270,$$

which are both larger than $\rho \sqrt{F(10,4;0.05)} = 0.123$, so the nonlinear behaviour of the model is likely to be much smaller than before. The angles between the hyperplanes are

$$w_1^N = 61.2°,$$

$$w_2^N = 59.9°.$$

The maximum intrinsic pl-curvatures associated with B_1 and B_2 are

$$(\Lambda_{max}^N)_1 = 0.25,$$

$$(\Lambda_{max}^N)_2 = 0.13,$$

and since $(\Lambda^N_{max})_1$ is exact

$$\Lambda^N_{max} = 0.25.$$

Also,

$$\Lambda^N_{RMS} = 0.08.$$

Thus, both intrinsic pl-curvature measures indicate that the tangent plane approximation to the solution locus is good.

The scalars α_{i1} and α_{i2} corresponding to B_1 and B_2 are

$$\alpha_{11} = 0.49, \qquad \alpha_{12} = 0.76,$$

$$\alpha_{21} = 0.47, \qquad \alpha_{22} = 0.69,$$

so that

$$v_1 = 0.49, \qquad w_1^c = 57.1°,$$

$$v_2 = 0.47, \qquad w_2^c = 55.7°.$$

Maximising the parameter effects pl-curvature associated with B_1 and B_2 over the angles ψ_{i1} and ψ_{i2} gives

$(\Lambda^T_{max})_1$ = 0.27 corresponding to ψ_{11} = 11.4°, ψ_{12} = 180.0°.

$(\Lambda^T_{max})_2$ = 0.13 corresponding to ψ_{21} = 9.2°, ψ_{22} = 180.0°.

Since $(\Lambda^T_{max})_1$ is exact, $\Lambda^T_{max} = 0.27$. Also, $\Lambda^T_{RMS} = 0.06$, so the uniform co-ordinates approximation is good.

The pl-curvatures indicate that the nonlinear behaviour of model (7.27) with the reduced data set is small in the region of interest and that linear approximation inferences can be made with security. For example 95% confidence intervals for $\theta_1, \theta_2, \theta_3, \theta_4$ obtained from a linear approximation are very similar to those obtained from a simulation (Table 7.3). Further, the simulated distributions of $\hat{\theta}_1$, $\hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$ are not skewed (Table 7.3) and the simulated distribution of $(\hat{\theta}_3 - \hat{\theta}_1 - 12.8(\hat{\theta}_2 - \hat{\theta}_4))$ shows no evidence of non-normality (Fig. 7.15).

7.7 NONLINEAR TRANSFORMATIONS OF θ

Often, $\theta = (\theta_1,...,\theta_p)'$ are not the parameters of interest. Instead, inference is required about parameters $\varphi = (\varphi_1,...,\varphi_p)'$, say, where φ is a nonlinear transformation of θ. For example, the parameters T_M and Ψ_M in the growth model (7.26) are of particular interest, since they represent the temperature at which maximum growth occurs and the maximum growth rate respectively. Thus, a confidence region for (ψ_M, T_M, b_1, b_2), say, is more useful than one for θ. This Section describes a method of assessing the validity of the uniform co-ordinates approximation to the φ parameter lines.

TABLE 7.3

Simulation results for model (7.27) - reduced data set

Parameter	Lin approx 95% conf. int.		Simulation 95% conf. int.	
θ_1	-0.33	-0.22	-0.33	-0.22
θ_2	0.074	0.088	0.074	0.087
θ_3	2.62	2.99	2.62	2.99
θ_4	-0.156	-0.136	-0.156	-0.136

Parameter	θ_1	θ_2	θ_3	θ_4
Skewness stat	0.47	-0.28	1.06	-1.03
Excess kurtosis stat	-1.76	-0.89	0.63	0.42

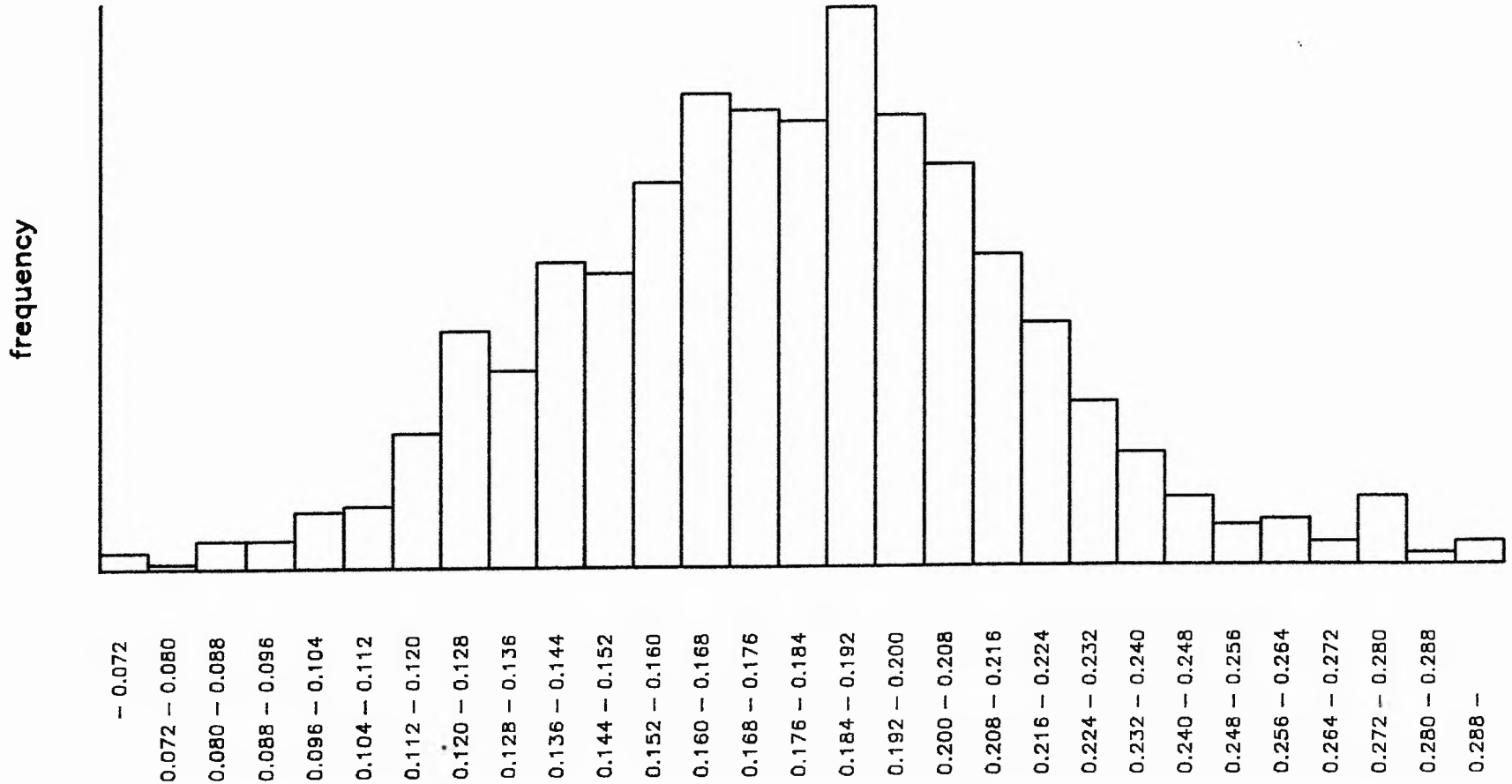


Fig. 7.15 The simulated distribution of $\hat{\theta}_3 - \hat{\theta}_1 - 12.8(\hat{\theta}_2 - \hat{\theta}_4)$: reduced data set.

Suppose that the Bates and Watts maximum parameter effects curvature Γ^T for the Ψ parameterisation is small. In this case, the Ψ parameter lines on sl(A) in the region of interest about $\hat{\eta}$ can be adequately approximated by straight, parallel equispaced lines; that is, by the parameter lines of a parameterisation ϕ, where ϕ is a linear transformation of Θ. If Λ^T_{max} is small, the uniform co-ordinates approximation to the ϕ parameter lines is good. Therefore, the Ψ parameter lines can be approximated by straight, parallel equi-spaced lines over the whole region of interest (assuming that the Ψ parameter lines on sl(B_1),...,sl(B_r) are also adequately approximated by the ϕ parameter lines). Thus, if Γ^T and Λ^T_{max} are both small, the uniform co-ordinates approximation to the Ψ parameter lines is likely to be good; if either of Γ^T and Λ^T_{max} is large, the approximation is likely to be poor.

For example, consider the growth model (7.26) and let

$$\Psi = (\psi_M,\ T_m,\ b_1,\ b_2)'.$$

When used in conjunction with the complete data set (ie including the data point corresponding to T = 13.6), the model has large intrinsic nonlinearity (Section 7.6.1), so that any linear approximation inference is likely to give a very misleading result (Section 3.2.2). In particular, a linear approximation should not be used to make inferences about the parameters of interest φ; instead, simulation based inferences could be used. (The large nonlinear behaviour is shown by a simulated distribution of $\hat{T}_M$, which has three clear modes (Fig. 7.16)). However, when the data point corresponding to T = 13.6 is removed, the intrinsic nonlinearity of the model is small (Section 7.6.4). Further, Λ^T_{max} = 0.27 and Γ^T = 0.24 so that, by the arguments above, the uniform co-ordinates approximation to the Ψ parameter lines is good.

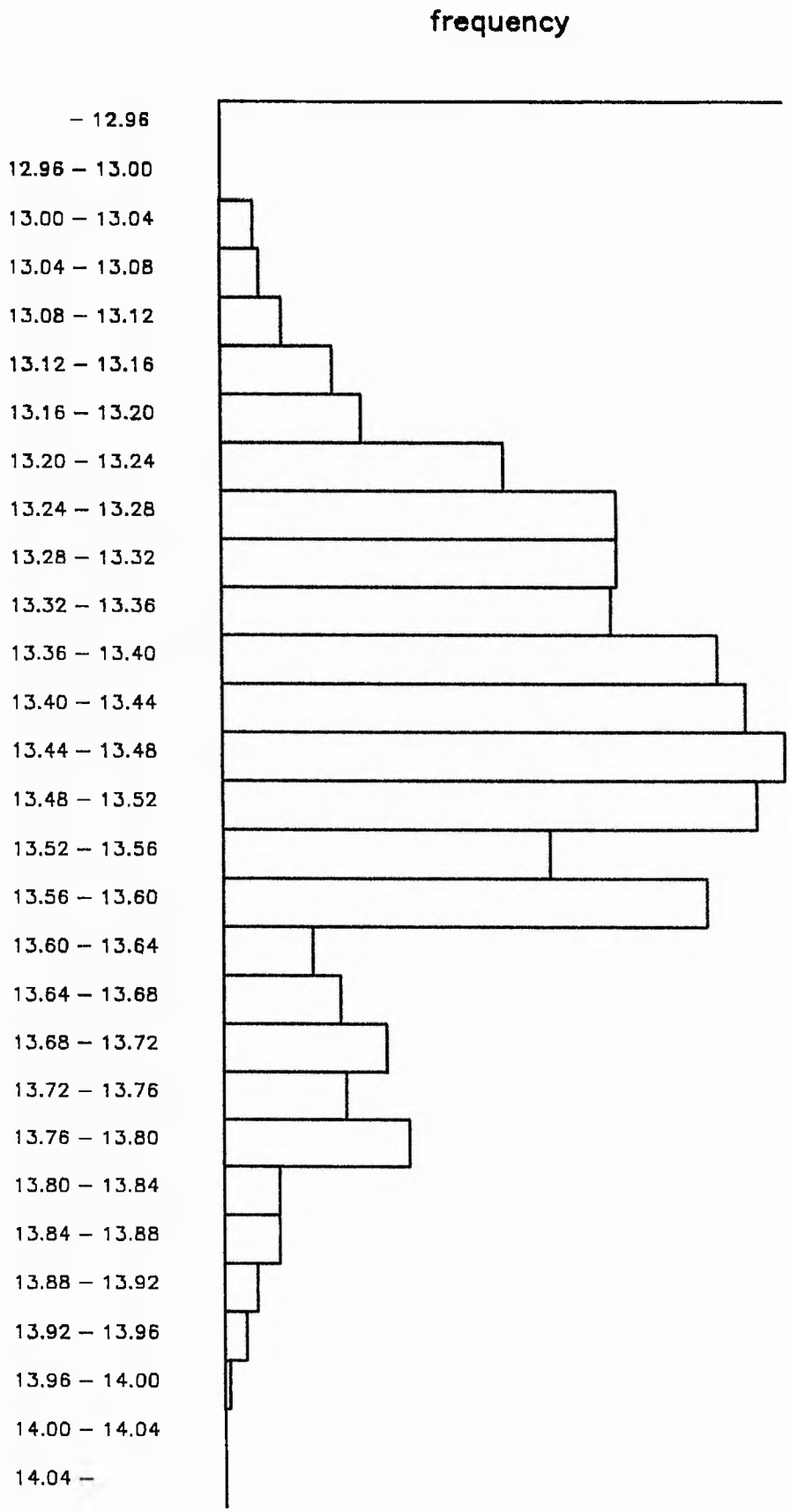


Fig. 7.16 The simulated distribution of $\hat{f}_m$.

Hence, linear approximation inferences can be made with security. For example, 95% confidence intervals for ψ_M, T_M, b_1, b_2 obtained from a linear approximation are close to those obtained from a simulation (Table 7.4). (A simulated distribution of $\hat{T}_M$ (reduced data set) is shown in Figure 7.17 for comparison with Figure 7.16; in the case of the reduced data set, the distribution gives no evidence of nonlinear behaviour).

7.8 DISCUSSION

The pl-curvatures have two major advantages. First, although the calculation of the exact maximum pl-curvatures and the RMS pl-curvatures is quite complicated (due to the construction of the L matrices and the need to determine which hyperplane is hit by a particular half lifted line), the calculation of the upper bounds $(\Lambda^N_{max})^*$ and $(\Lambda^T_{max})^*$ is very straightforward. For example, $(\Lambda^N_{max})^*$ is obtained by calculating $(\Lambda^N_{max})_i$ for each hyperplane B_i, $i = 1...r$; this only requires the knowledge of e_{i1}, e_{iB} and ξ_i, $i = 1...r$, which can be found by linear regression using any standard statistical package (eg GENSTAT 5 (Payne *et al.*, 1987)). Similarly, $(\Lambda^T_{max})^*$ is obtained by finding α_{i1}, α_{i2}, $i = 1...r$, again by linear regression, and then maximising functions of two variables (which can also be done in GENSTAT5). At no stage is it necessary to calculate the L matrices or the orthonormal basis E_1 (other than e_{11} and e_{12}). Usually $(\Lambda^N_{max})^*$ and $(\Lambda^T_{max})^*$ are equal to Λ^N_{max} and Λ^T_{max} respectively; even if this is not the case, the upper bounds are very useful measures of intrinsic and parameter effects nonlinearity and are generally sufficient to assess the nonlinearity of a model. It is helpful to be able to compute and compare the exact maximum and RMS pl-curvatures; however, in many cases, the extra

TABLE 7.4

Simulation results for model (7.26) - reduced data set

Parameter	Lin approx 95% conf. int.		Simulation 95% conf. int.	
ψ_M	0.786	0.857	0.784	0.854
T_M	13.27	13.90	13.27	13.93
b_1	0.074	0.088	0.074	0.087
b_2	-0.156	-0.136	-0.156	-0.136

Parameter	ψ_M	T_M	b_1	b_2
Skewness stat	-0.32	1.05	-0.28	-1.03
Excess kurtosis stat	-0.47	1.64	-0.89	0.42

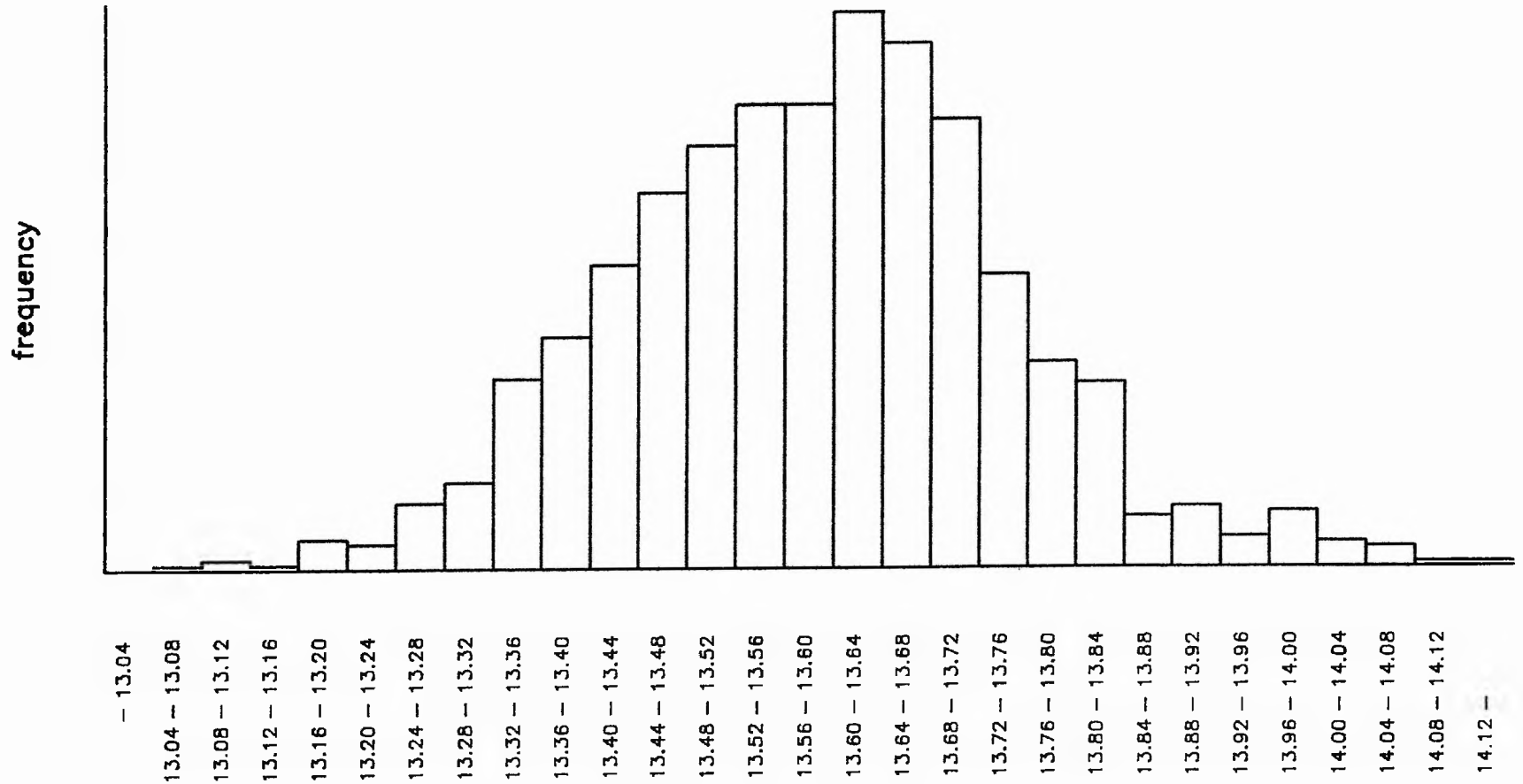


Fig. 7.17 The simulated distribution of $\hat{f}_m$: reduced data set.

computation gives only limited extra information about the nonlinear behaviour of the model.

The second advantage is that the pl-curvatures can identify severe nonlinear behaviour outside the region of interest defined by the linear approximation. For example, suppose that $n = 2$, $p = 1$ and the solution locus can be represented by a pair of line segments sl(A) and sl(B) with an angle $w^N = \pi/2$ between them (Fig. 7.18). Suppose that y, ξ and $\rho\sqrt{F}$ are as shown in the Figure and that σ^2 is estimated independently by s^2. The linear approximation region of interest is the line segment RS and the tangent plane approximation is exact within this region. However, if ξ is not much larger than $\rho\sqrt{F}$, inference using the tangent plane approximation is poor. For example, a linear approximation 95% confidence interval for θ consists of those values of θ for which $\eta(\theta)$ lies in RS, whereas an exact 95% interval consists of those values for which $\eta(\theta)$ lies in $RS \cup TZ$. The intrinsic pl-curvature in direction e_1 is

$$\Lambda^N_{e_1} = \frac{\rho \sin(w^N)}{\xi(1 + \cos(w^N))} = \frac{\rho}{\xi}$$

which is large compared to $1/\sqrt{F}$ indicating the severe nonlinear behaviour.

A disadvantage of the pl-curvatures is that, because they are conservative measures (Section 7.4.2.3), they can indicate severe nonlinear behaviour when it does not actually exist (although this is preferable to failing to indicate severe nonlinearity when it does exist). For example, suppose that $n = 2$, $p = 1$ and the solution locus consists of two line segments sl(A) and sl(B) as before, but that w^N and ξ are very

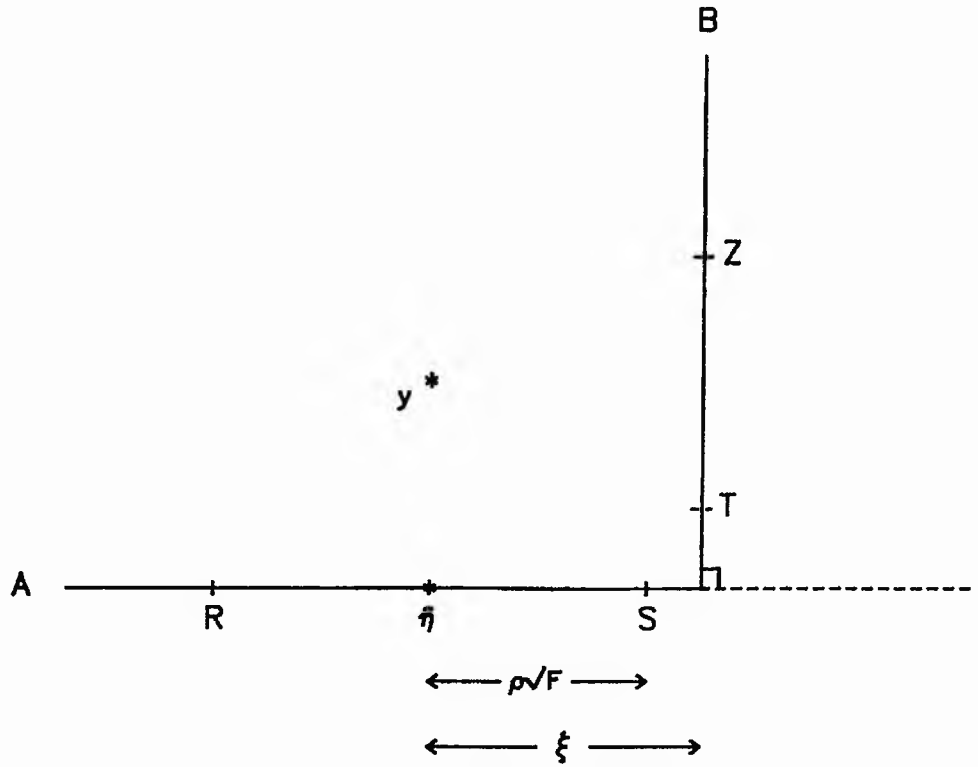


Fig. 7.18 Large nonlinear behaviour outside the linear approximation region of interest.

small compared to $\rho\sqrt{F}$ (Fig. 7.19). The maximum deviation of the solution locus from the tangent plane in the linear region of interest is $(\rho\sqrt{F} - \xi)\tan(w^N)$, which is small, but the intrinsic pl-curvature can be made arbitrarily large by taking ξ arbitrarily small. However, problems of this type can be overcome by examining the values of ξ and w^N when calculating $\Lambda_{e_i}^N$; if $\Lambda_{e_i}^N$ is high, but ξ is small, the maximum deviation can be calculated to determine whether the intrinsic nonlinearity is overestimated by the intrinsic pl-curvature.

The pl-curvatures are constructed by assuming that the region of interest of the solution locus consists of sl(A) and the sl-segments of the hyperplanes that border sl(A), B_1,..., B_r. In general, a half lifted line consists of a large number of line segments on a large number of hyperplanes; however, all but the "first two" line segments are assumed to be sufficiently far away from $\hat{\eta}$ that they can be ignored. In the growth model example, this is a reasonable assumption since the only changes of hyperplane close to the region of interest are those corresponding to the lines T_M = 12.8 and T_M = 13.6; once a half lifted line has hit its first hyperplane (ie at T_M = 12.8 or T_M = 13.6), it does not hit another until it is far from the region of interest (ie at T_M = 10.8 or T_M = 15.0). If the assumption is not valid because there are many changes of hyperplane near the region of interest, the pl-curvatures assess the nonlinearity due to the hyperplanes bordering sl(A), but are of only limited use in assessing the overall nonlinearity of the model. In general, the more hyperplanes near the region of interest, the greater the nonlinear behaviour; inferences based on linear approximations to models with many changes of hyperplane near the region of interest should be treated with caution.

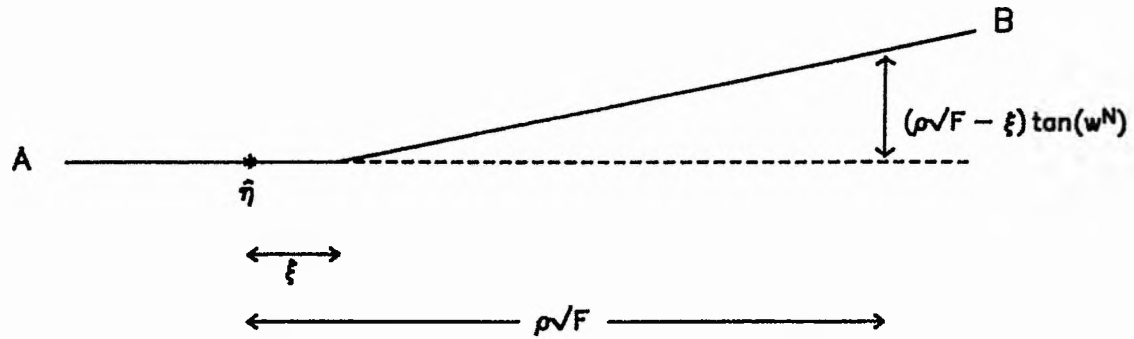


Fig. 7.19 Large pl–curvature when ξ is small.

Finally, since the nonlinear behaviour of piecewise linear change-point models increases with the number of changes of hyperplane in the region of interest, there is a conflict of interests when fitting and making inference from such models. In selecting an appropriate model to describe the relationship between two variables, it is important to have a full spread of experimental design settings. If the relationship changes more rapidly in a particular area (ie near a change point), then it is also important to have a number of design points in this area. For example, the growth rates at temperatures close to 13.4°C are necessary to determine whether the growth rate model is piecewise linear with a change point or whether the growth rate changes smoothly with temperature in the region of maximum growth (cf Section 8.3). However, it is the data points near the change points of a piecewise linear model that cause large nonlinear behaviour and hence problems of inference. For example, once the piecewise linear growth model has been selected, inference is made much easier if the data points near $T = 13.4$ are removed. Intuitively, inferences based on a linear approximation to a piecewise linear change-point model with all the data points near the change points removed might be expected to be easier to calculate and to present but also more conservative than "exact" inferences based on the complete data set; further study is required to investigate this subject in detail.

CHAPTER 8 THE GROWTH OF BROWN TROUT FED ON MAXIMUM RATIONS

8.1 INTRODUCTION

Elliott (1975) describes an experiment to investigate the growth of brown trout fed on maximum rations. Elliott shows that the growth rate of a trout depends both on the weight of the trout and on the water temperature. Further, he develops a growth model that describes the relationship between the weight of a trout at time t' and

i) the weight of the trout at some previous time t,

ii) the water temperature T,

given that the trout is fed on maximum rations between t and t'.

A growth model is required in the development of a model of survival between t_1 (the May/June electro-fishing) and t_2 (the August/September electro-fishing) in Black Brows Beck (Chapter 9). Although Elliott's growth model could be used for this purpose, the model suffers from a number of limitations (which are described in Section 8.3.1). Hence, in this Chapter, the results of Elliott's experiment are reanalysed and a new growth model is developed.

First, the experiment is briefly summarised. The scope of the Chapter is then outlined.

8.1.1 The Growth Experiment

The growth experiment consisted of fifteen smaller experiments. In each of these "sub-experiments", hatchery brown trout, in good physical condition, were placed in tanks and were fed to appetite satiation for a period of usually 35 or 42 days. Water was supplied to the tanks from Windermere and was at an approximately constant temperature throughout any one sub-experiment; thus, each sub-experiment corresponded to an investigation of maximum growth at a particular temperature. The temperatures varied between sub-experiments from 3.8°C to 21.7°C.

The trout were from four weight classes; namely, close to 10 g, 50 g, 90 g and 250 g. In each sub-experiment, one of the following arrangements of trout were adopted:

i) one 50-g trout (six sub-experiments),

ii) four 50-g trout (four sub-experiments),

iii) four 50-g trout and one trout each of 10 g, 90 g and 250 g (five sub-experiments).

Each fish was weighed at the start and end of the sub-experiment. In addition, the 50-g trout in ii) and iii) above were weighed twice more, generally after 14 and 28 days. Two 50-g trout died during the sub-experiment at the highest temperature (21.7°C); therefore, this sub-experiment was stopped after only 28 days. The experimental design is summarised in Table 8.1. A table of results is given in Elliott (1975).

Two tanks were used in each sub-experiment. Each tank was divided into four compartments and the trout in i), ii) or iii) above were each allocated to one of the

TABLE 8.1

A summary of the experimental design

Temperature / duration of expt		Weight class			
		10 g	50 g	90 g	250 g
3.8°C / 42 days	no of trout	-	4	-	-
	no of weighings	-	4	-	-
4.2°C / 42 days	no of trout	-	1	-	-
	no of weighings	-	2	-	-
5.6°C / 42 days	no of trout	1	4	1	1
	no of weighings	2	4	2	2
6.8°C / 42 days	no of trout	-	1	-	-
	no of weighings	-	2	-	-
7.1°C / 42 days	no of trout	-	4	-	-
	no of weighings	-	4	-	-
9.5°C / 35 days	no of trout	1	4	1	1
	no of weighings	2	4	2	2
10.8°C / 35 days	no of trout	-	1	-	-
	no of weighings	-	2	-	-
12.8°C / 42 days	no of trout	1	4	1	1
	no of weighings	2	4	2	2
13.6°C / 42 days	no of trout	-	1	-	-
	no of weighings	-	2	-	-
15.0°C / 35 days	no of trout	1	4	1	1
	no of weighings	2	4	2	2
16.2°C / 35 days	no of trout	-	1	-	-
	no of weighings	-	2	-	-
17.8°C / 42 days	no of trout	-	4	-	-
	no of weighings	-	4	-	-
19.5°C / 35 days	no of trout	1	4	1	1
	no of weighings	2	4	2	2
20.4°C / 35 days	no of trout	-	1	-	-
	no of weighings	-	2	-	-
21.7°C / 28 days	no of trout	-	4+	-	-
	no of weighings	-	3	-	-

+ two of these trout died during the experiment

eight compartments at random. (The remaining compartments were then filled with trout that were used for quite separate feeding experiments). It is important to note that both tanks were supplied by water from the same source, so that the growth of the trout within a sub-experiment were not independent (since slight fluctuations in water temperature, water content, etc were the same for all the fish). In fact, the dependence was such that the between and within tank variation in the growth of the replicate 50-g trout was very small and yielded negligible information about the "true" between fish variation in growth (see Section 8.3.1). Hence, the allocation of trout to tanks and to compartments is not considered.

8.1.2 The Scope of this Chapter

The objective of this Chapter is to develop a model of the growth in weight of brown trout fed on maximum rations. However, it is convenient to analyse first the growth rate of brown trout fed on maximum rations. Section 8.2 defines the growth rate of an individual trout and hence the variable $\bar{g}$, the average of the growth rates of the trout from the same weight class kept at the same water temperature. In Sections 8.3 and 8.4, the relationship between $\bar{g}$ and fish weight and water temperature is investigated. Since most of the available information concerns 50-g trout, a model is first developed that describes the relationship between $\bar{g}$ and water temperature for trout from this weight class. The model is then extended to include trout from the 10 g, 90 g and 250 g weight classes. Finally, in Section 8.5, the model of growth rates is used to develop a model of the growth in weight of brown trout fed on maximum rations.

8.2 GROWTH RATES

Plots of log fish weight against time (Fig. 8.1) suggest that, over the comparatively short time period of a sub-experiment, the growth of a trout can be well described by the model

$$\log(w_i) = \alpha + \frac{g}{100} t_i + \epsilon_i, \qquad i = 1 \ldots m, \tag{8.1}$$

where

i) the trout is weighed at times t_i, $i = 1 \ldots m$,

ii) w_i is the weight of the trout at time t_i, $i = 1 \ldots m$,

iii) the errors $\{\epsilon_i\}$ are NID$(0,\tau^2)$,

iv) the parameters a, g and τ^2 depend on the individual fish.

The parameter g is interpreted as the growth rate of the trout or, in Elliott's notation, the mean specific increase in live weight.

Let weight class 1,2,3,4 correspond to trout of initial weight 10 g, 50 g, 90 g, 250 g respectively. Also, let n_{ij} be the number of trout in weight class i at temperature T_j, $i = 1 \ldots 4$, $j = 1 \ldots 15$. Then, the growth rate g_{ijk} of the kth fish in weight class i at temperature T_j is estimated from (8.1) by linear least squares regression, $i = 1 \ldots 4$, $j = 1 \ldots 15$, $k = 1 \ldots n_{ij}$. In all cases, the fit of model (8.1) is excellent and examinations of the residuals give no evidence that the form of the error term is inappropriate.

Elliott (1975) analyses the relationship between $\{g_{ijk}\}$ and fish weight and water temperature. However, due to the non-independence of the growth rates of the

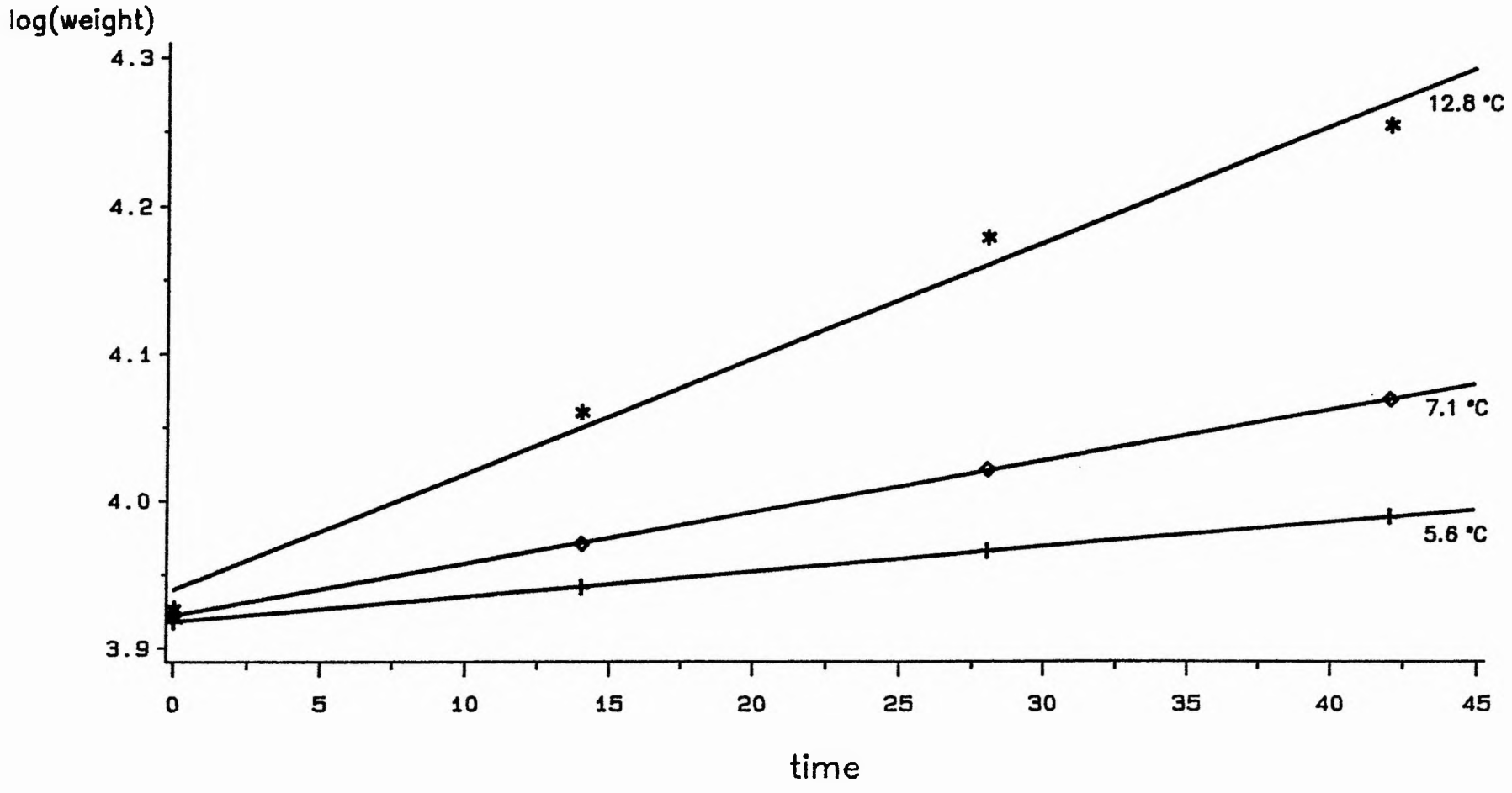


Fig. 8.1 The growth of 50−g trout.

fish in the same sub-experiment, the variable analysed in this Chapter is $\bar{g}$, where $\bar{g}_{ij}$ is the average of the growth rates of the fish in weight class i at temperature T_j

$$\bar{g}_{ij} = \begin{cases} \sum_{k=1}^{n_{ij}} g_{ijk}/n_{ij} & \text{if } n_{ij} > 0 \\ \text{undefined} & \text{if } n_{ij} = 0 \end{cases}, \qquad i = 1...4, \quad j = 1...15.$$

It is noted that the variation in the growth rates of the 50-g trout within any sub-experiment was always very small, except at the highest temperature. In this sub-experiment, the two trout that survived had much higher growth rates than the two that died. Since the dead fish showed growth patterns divergent from those of the other fish, the results obtained from the two dead fish are omitted and $\bar{g}_{2,15}$ is the average of the growth rates of the two fish that survived. The values of $\{\bar{g}_{ij}\}$ are given in Table 8.2.

8.3 THE RELATIONSHIP BETWEEN GROWTH RATE AND TEMPERATURE FOR 50-g TROUT

This Section investigates the relationship between growth rate and temperature for 50-g trout. Elliott's model of individual growth rates $\{g_{2jk}\}$ is described first. A new model relating $\{\bar{g}_{2j}\}$ to water temperature is then developed.

8.3.1 The Elliott Model

Elliott divides the temperature range empirically into four intervals (3.8-12.8, 12.8-13.6, 13.6-19.5, 19.5-21.7°C), in each of which, the growth rates $\{g_{2jk}\}$ are found to vary linearly with temperature (Fig. 8.2). In each interval it is assumed that

TABLE 8.2

Values of $(\bar{g})$

Temperature	Weight class			
	10 g	50 g	90 g	250 g
3.8°C	-	0.020	-	-
4.2°C	-	0.062	-	-
5.6°C	0.279	0.169	0.155	0.101
6.8°C	-	0.254	-	-
7.1°C	-	0.351	-	-
9.5°C	0.799	0.499	0.390	0.300
10.8°C	-	0.592	-	-
12.8°C	1.190	0.747	0.630	0.476
13.6°C	-	0.743	-	-
15.0°C	0.990	0.616	0.520	0.400
16.2°C	-	0.418	-	-
17.8°C	-	0.216	-	-
19.5°C	0.014	0.010	0.008	0.006
20.4°C	-	-0.217	-	-
21.7°C	-	-0.367	-	-

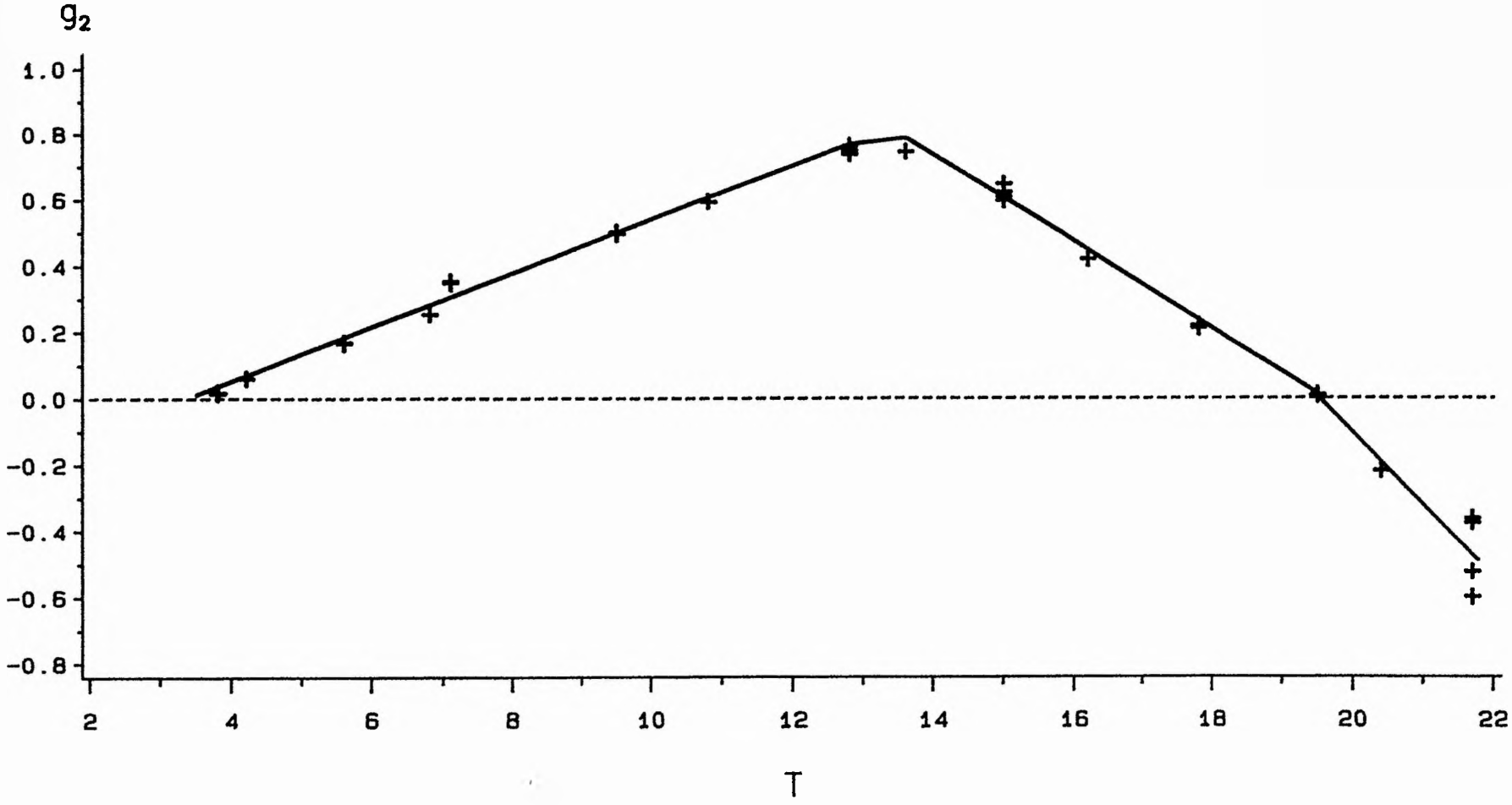


Fig. 8.2 Elliott's model of the growth rate of 50–g trout.

$$g_{2jk} = a_1 + a_2 T_j + \epsilon_{jk}, \qquad (8.2)$$

where $\{\epsilon_{jk}\}$ are NID$(0, \sigma^2)$. The parameters a_1, a_2 and σ^2 corresponding to the first, third and fourth intervals are estimated by linear least squares regression. The parameters a_1 and a_2 corresponding to the second interval are estimated so that the expected growth rate is a continuous function of temperature (ie so that the line segment in the second interval "joins up" the line segments in the first and third intervals).

Elliott's model closely follows the changes in growth rate with temperature (Fig. 8.2). However, the model has a number of unsatisfactory features. First, the model has a large number of parameters; there are two parameters for each of the straight lines in the first, third and fourth intervals; further, the temperatures 12.8, 13.6 and 19.5°C, which define the four temperature intervals, effectively constitute three change-point parameters which are estimated by eye. Secondly, the model lacks biological realism; although the model describes the data very well empirically, it is difficult to envisage biological processes that lead to the sharp changes in slope at the temperatures 12.8, 13.6 and 19.5°C. Thirdly, Elliott's model assumes that all the individual growth rates are independent and (within each temperature interval) gives equal weight to each observation. However, a plot of the residuals obtained when model (8.2) is fitted to the data for the first temperature interval indicates that the growth rates of the replicate 50-g trout are not independent (Fig. 8.3) (cf Section 8.1.1).

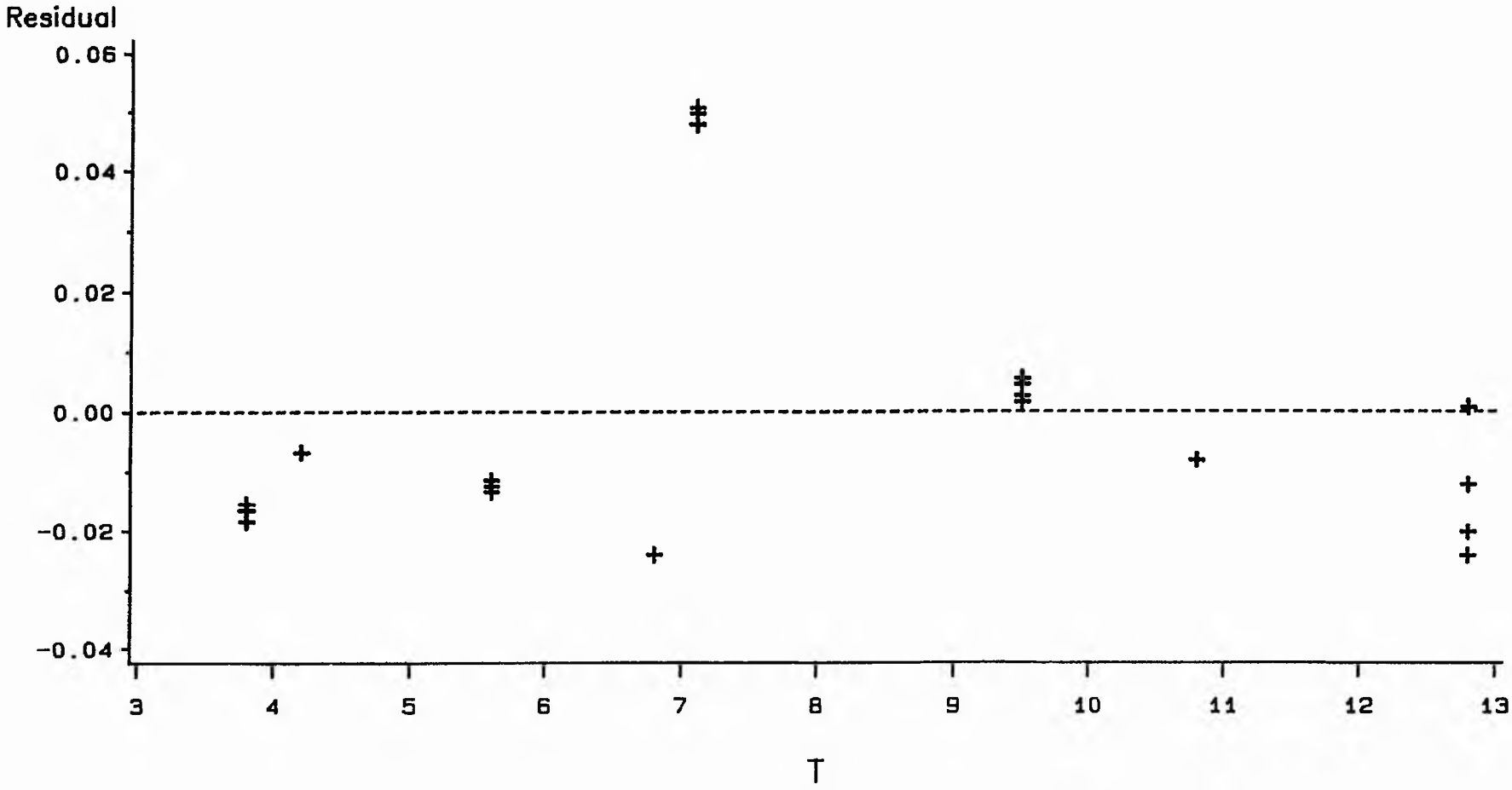


Fig. 8.3 Residuals from Elliott's model showing the non–independence of the growth rates of the replicate 50–g trout.

8.3.2 The Relationship Between $\bar{g}_2$ and T: Pair-of-Straight-Lines Model

An inspection of a graph of $\{\bar{g}_{2j}\}$ against $\{T_j\}$ (Fig. 8.4) suggests that the relationship between $\bar{g}_2$ and T might be well described by a pair-of-straight-lines model in which the temperature range is divided into just two intervals, in each of which $\bar{g}_2$ varies linearly with temperature. This gives a four parameter model

$$\bar{g}_{2j} = \begin{cases} \psi + b_1(T_j - T_M) + \epsilon_j & \text{if } T_j \leq T_M \\ \psi + b_2(T_j - T_M) + \epsilon_j & \text{if } T_j > T_M \end{cases}, \qquad j = 1...15, \qquad (8.3)$$

where $\{\epsilon_j\}$ are independent and where $\epsilon_j \sim N(0, \sigma^2/v_j)$, $j = 1...15$, for weights $\{v_j\}$. The reduction from four temperature intervals in Elliott's model to just two in model (8.3) is plausible because the omission of the two dead fish increases the average growth rate at 21.7°C and because the second of Elliott's temperature intervals is so short. In (8.3)

i) T_M is interpreted as the temperature at which maximum growth occurs,

ii) ψ is interpreted as the maximum growth rate,

iii) b_1, b_2 are the gradients of the two straight lines.

Model (8.3) is fitted by weighted nonlinear least squares regression. The choice of suitable weights $\{v_j\}$ depends on a number of factors, such as

i) the number of times each fish was weighed (some 50-g trout were weighed four times, whilst others were weighed only twice (Table 8.1)),

ii) the number of fish used to calculate $\bar{g}_2$,

iii) the variation in growth rate might be expected to increase as the growth rate increases and as the water temperature increases.

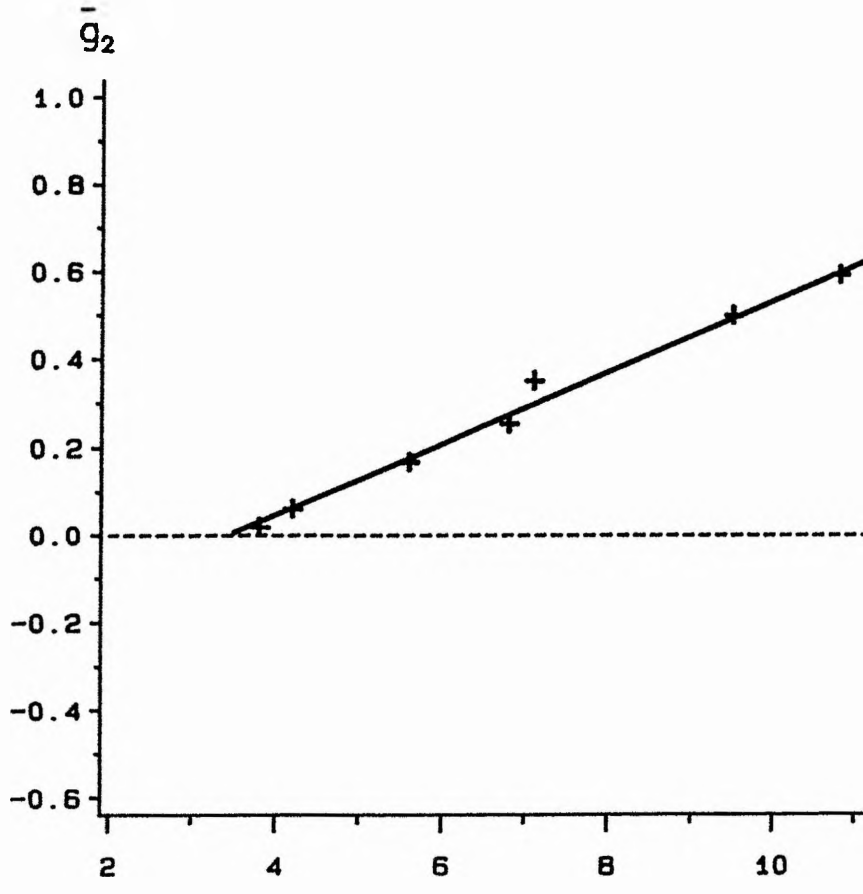
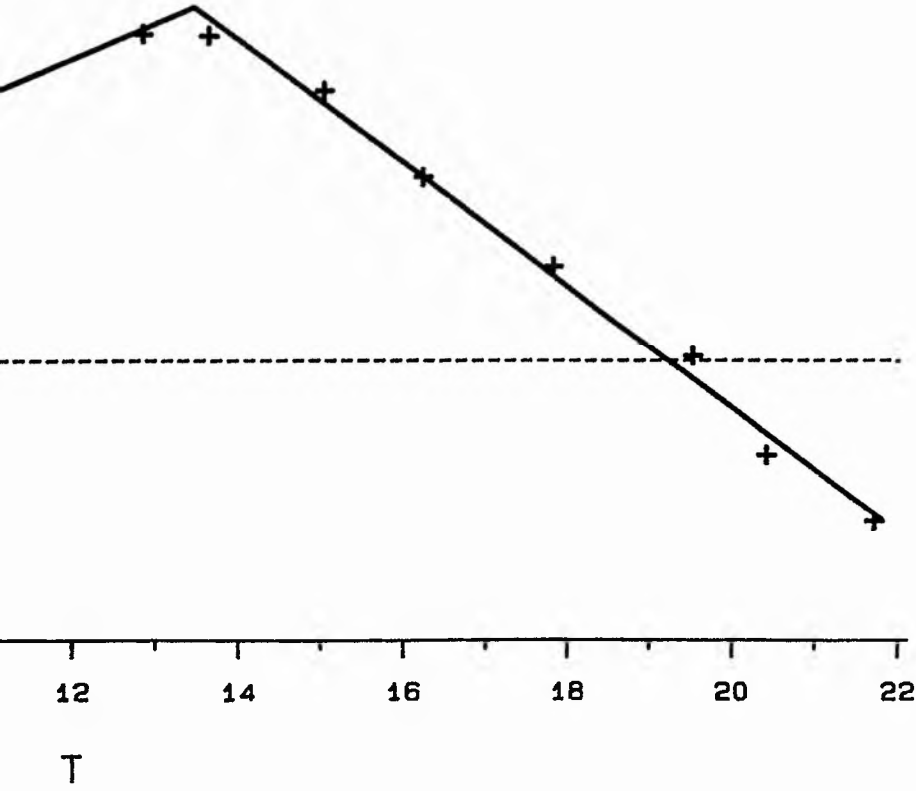


Fig. 8.4 The pair—of-



-straight-lines model (8.4).

First, consider i) above. In the notation of equation (8.1), the standard error of the estimate of the growth rate of an individual fish is

$$100\tau/\sqrt{\{\sum_{i=1}^{m}(t_i-\bar{t})^2\}}.$$

Thus, the growth rate of a fish is generally estimated more precisely the more times the fish is weighed. Therefore, it is arguable that values of $\bar{g}_2$ obtained from fish weighed four times merit more weight in the analysis than values of $\tilde{g}_2$ obtained from fish weighed only twice. However, the standard errors of the growth rate estimates are always very small, regardless of the number of times the fish is weighed. (Where a trout is weighed only twice, the variance τ^2 can not be estimated from the data, but is assumed to be of a similar magnitude to those of the other trout). In particular, the standard errors are always very small compared to the between fish variation in growth rates (as, for example, estimated by σ^2 in Elliott's model (8.2)). Hence, weighing a trout four times rather than two results in only a small increase in information and the effect of i) can be ignored.

Similarly, in ii), it is arguable that values of $\bar{g}_2$ based on the growth rates of four fish will be more precisely known than values of $\bar{g}_2$ based on the growth rate of only one fish. However, the increase in information is much smaller than the number of fish due to the non-independence of the individual growth rates (cf Fig. 8.3). Since the net effect of ii) and iii) is unknown, the simplest initial approach is to give a uniform weighting to each point.

The model

$$\bar{g}_{2j} = \begin{cases} \psi + b_1(T_j - T_M) + \epsilon_j & \text{if } T_j \le T_M \\ \psi + b_2(T_j - T_M) + \epsilon_j & \text{if } T_j > T_M \end{cases}, \qquad j = 1 \ldots 15, \qquad (8.4)$$

where $\{\epsilon_j\}$ are NID$(0, \sigma^2)$, is fitted by (unweighted) nonlinear least squares regression and is an excellent description of the data (Fig. 8.4). Further, an examination of the residuals gives no evidence that the assumption of constant variance is inappropriate. Parameter estimates are given in Table 8.3. Since, model (8.4) is a change-point model, the nonlinear behaviour of the model can not be assessed by the standard measures of nonlinearity (Chapter 7). However, with a suitable reparameterisation, model (8.4) can be written as a piecewise linear change-point model (Section 7.6), so the pl-curvature measures of Chapter 7 can be applied. The intrinsic nonlinearity of the model is large ($\Lambda^N_{max} = 0.91$, $1/(2\sqrt{F(4,11;0.05)}) = 0.27$), so linear approximation inferences could give very misleading results. Hence, simulation parameter 95% confidence limits are given in Table 8.3.

It is interesting to note that if model (8.3) is fitted with weights (ν_j = number of 50-g trout at temperature T_j), the parameter estimates and confidence limits are virtually unchanged. However, since the error structure of model (8.4) appears adequate, then for simplicity, it is concluded that (8.4) is a good description of the relationship between $\bar{g}$ and temperature for 50-g trout.

It is convenient to note at this stage that model (8.4) can be reparameterised as

TABLE 8.3

Parameter estimates for models (8.4) and (8.5)

Parameter	Estimate	95% confidence limits	
ψ	0.809	0.771	0.843
T_M	13.42	13.12	13.82
b_1	0.0807	0.0729	0.0885
b_2	-0.140	-0.151	-0.131
T_L	3.40	2.89	3.84
T_U	19.19	19.00	19.41

rss = 0.0119 on 11 d.f.
s^2 = 0.00108

$$\bar{g}_{2j} = \begin{cases} \psi \dfrac{T_j - T_L}{T_M - T_L} + \epsilon_j & \text{if } T_j \le T_M \\ \psi \dfrac{T_j - T_U}{T_M - T_U} + \epsilon_j & \text{if } T_j > T_M \end{cases}, \qquad j = 1 \ldots 15, \quad (8.5)$$

where $\{\epsilon_j\}$ are NID$(0, \sigma^2)$. In (8.5), the parameters T_L and T_U have the useful biological interpretations of the lower and upper temperatures at which zero growth occurs. Parameter estimates and simulation parameter 95% confidence limits for T_L and T_U are also given in Table 8.3.

8.3.3 The Relationship Between $\bar{g}_2$ and T: Alternative Models

Model (8.4) is an excellent description of the relationship between $\bar{g}$ and T for 50-g fish. Further, the model overcomes many of the limitations of Elliott's model since

i) model (8.4) has only a small number of parameters (ie only four parameters),

ii) the number of temperature intervals is reduced from four to two,

iii) there is no evidence that the error structure of (8.4) is inappropriate.

However, the sharp change in physiological reaction implied by the instantaneous change in slope at $T = T_M$ does seem biologically implausible. Therefore, extended efforts have been made to fit alternative models to the data for 50-g fish. The aim was to find a model that fitted the data as well as model (8.4) but which had a smoother transition between the ascending and descending parts of the curve. However, as might be expected, most of the models investigated could not accommodate the sharp peak in the region of maximum growth and consistently underestimated $\bar{g}_2$ in this region and overestimated $\bar{g}_2$ at low and high temperatures.

The one model found that suggested an improvement was the five parameter hyperbola

$$\bar{g}_{2j} = \psi + \frac{1}{2}\{(b_1 + b_2)(T_j - T_M) - (b_1 - b_2)((T_j - T_M)^2 + c)^{\frac{1}{2}}\} + \epsilon_j,$$

$$j = 1 \ldots 15. \qquad (8.6)$$

where $\{\varepsilon_j\}$ are NID$(0,\sigma^2)$. The hyperbola is a generalisation of the pair-of-straight-lines model (8.4), in which

i) (T_M, ψ) are the temperature and growth rate co-ordinates of the vertex of the axes of the hyperbola,

ii) b_1, b_2 are the gradients of the axes of the hyperbola,

iii) $c \geq 0$ measures the "curvature" of the hyperbola near the vertex; when $c = 0$, model (8.6) reduces to model (8.4).

The hyperbola (8.6) is an excellent description of the data (Fig. 8.5). An examination of the residuals gives no evidence that the form of the error term is inappropriate. Parameter estimates are given in Table 8.4. The intrinsic nonlinearity of the model is large ($\Gamma^N = 0.48$, $1/(2\sqrt{F(5,10;0.05)}) = 0.27$), so simulation based confidence limits are given in Table 8.4.

The hyperbola (8.6) is more realistic biologically than the pair-of-straight-lines (8.4) because it is smooth in the region of maximum growth. However, it is also a more complex model, since it has an extra parameter. Therefore, the use of the hyperbola to model the observed data is only justified if it explains features of the data that are unaccounted for by model (8.4). At first sight, this appears to be the case, since the estimate of the parameter c is significantly greater than zero (F =

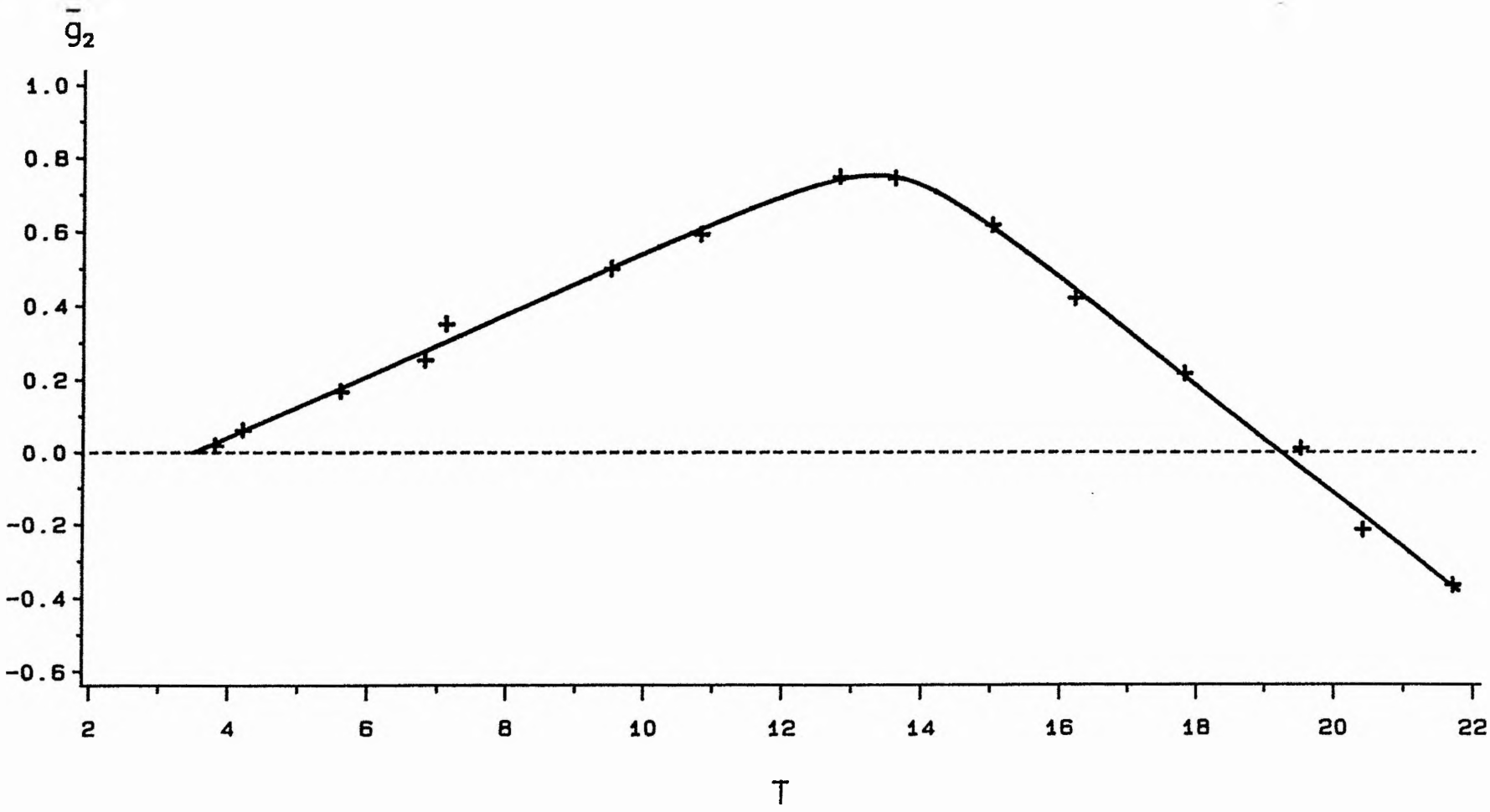


Fig. 8.5 The hyperbola model (8.6).

TABLE 8.4

Parameter estimates for model (8.6)

Parameter	Estimate	95% confidence limits	
ψ	0.858	0.801	0.961
T_M	13.55	13.19	13.89
b_1	0.0847	0.0763	0.0964
b_2	-0.150	-0.166	-0.137
c	0.894	0.000	3.609

rss = 0.00835 on 10 d.f.
s^2 = 0.000835

4.21, F(1,10;0.05) = 4.96, F_{sim} (0.05) = 2.80, p < 0.05). However, evidence for the adequacy of the model with $c = 0$ is presented in Section 8.4. For the time being, both (8.4) and (8.6) are considered as suitable models of the relationship between $\bar{g}_2$ and T.

8.4 THE RELATIONSHIP BETWEEN $\bar{g}$ AND T: ALL WEIGHT CLASSES

The relationships between $\bar{g}$ and temperature for each of the other three weight classes are also well described by the pair-of-straight-lines model (8.4). (This is to be expected, since there are only five data points corresponding to each of these weight classes). This suggests formulating a "global" pair-of-straight-lines model

$$\bar{g}_{ij} = \begin{cases} \psi_i + b_{1i}(T_j - T_{Mi}) + \epsilon_{ij} & \text{if } T_j \le T_{Mi} \\ \psi_i + b_{2i}(T_j - T_{Mi}) + \epsilon_{ij} & \text{if } T_j > T_{Mi} \end{cases}, \qquad i = 1...4, \quad j = 1...15, \qquad (8.7)$$

where $\{\epsilon_{ij}\}$ are NID$(0,\sigma^2)$. In (8.7), the relationship between g and temperature for each weight class is described by a pair-of-straight-lines model. The model has 16 parameters, four for each weight class, where, for example, T_{Mi} is the temperature at which maximum growth occurs for fish in the ith weight class, i = 1...4.

Model (8.7) is fitted by nonlinear least squares regression and is an excellent description of the data (Fig. 8.6). An examination of the residuals reveals no evidence of heterogeneous variances. Further, although values of $\bar{g}$ corresponding to the same temperature are not strictly independent (because they were obtained from the same sub-experiment), the residuals reveal no strong pattern of dependence.

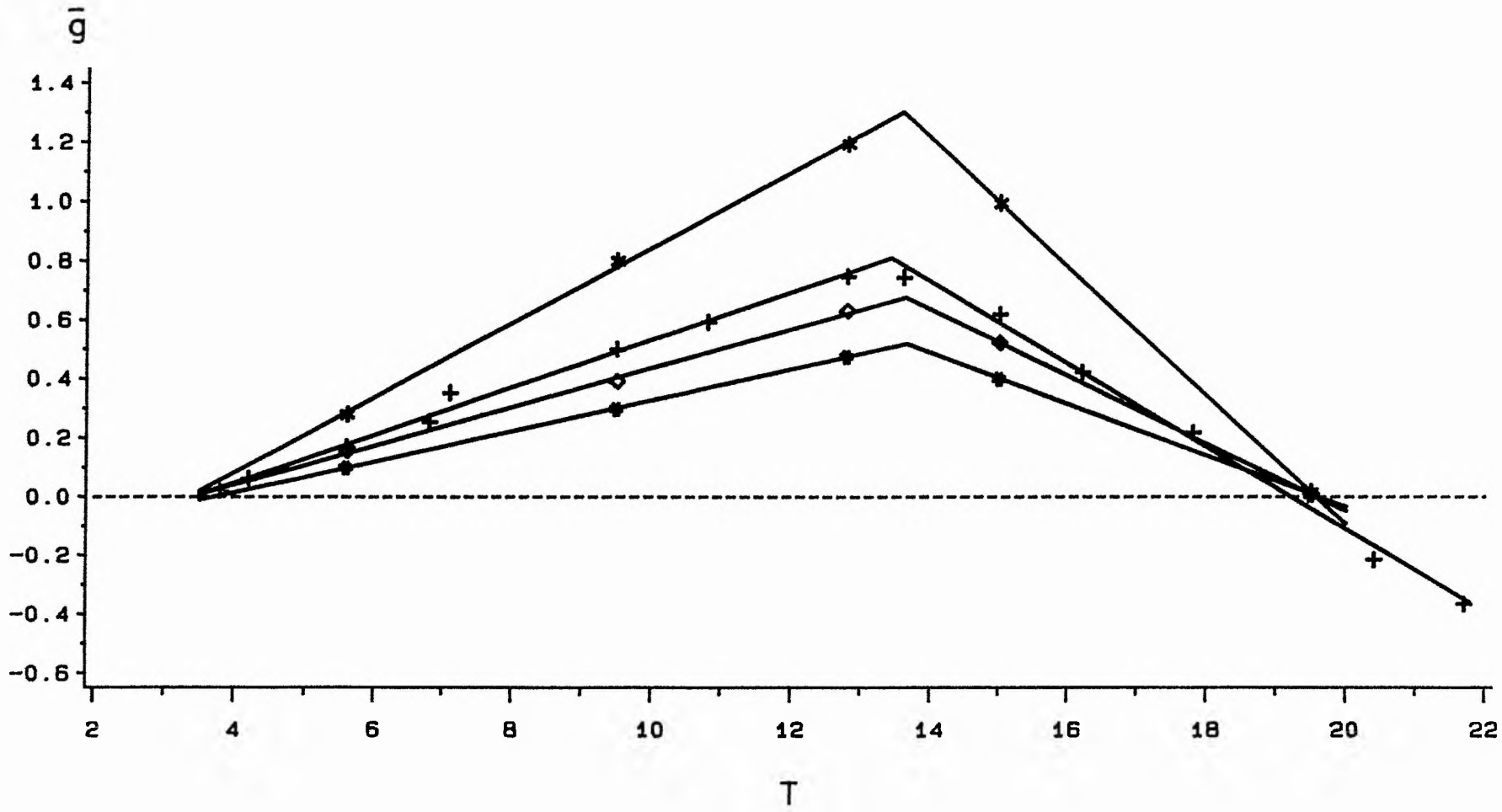


Fig. 8.6 The global pair–of–straight–lines model (8.7).

Hence, the error distribution in (8.7) is assumed to be adequate. Parameter estimates are given in Table 8.5. Model (8.7) can be written in the form of a piecewise linear change-point model with four change points corresponding to the temperatures $\{T_{Mi}\}$. Hence, the pl-curvatures of Chapter 7 are used to assess the nonlinear behaviour of the model. The intrinsic nonlinearity is very large (Λ^N_{max} = 1.67, $1/(2\sqrt{F(16,14;0.05)})$ = 0.32), so linear approximation inferences can not be used. Hence, simulation based confidence limits are given in Table 8.5.

Model (8.7) contains too many parameters to be biologically useful. Hence, the relationship between the parameters across weight classes is now investigated. First, an inspection of Fig. 8.6 and Table 8.5 suggests that

i) the temperature at which maximum growth occurs might be the same for all weight classes,

ii) the lower and upper temperatures at which zero growth occurs might be the same for all weight classes.

This is most easily investigated by reparameterising (8.7) in the form of (8.5) as

$$\bar{g}_{ij} = \begin{cases} \psi_i \dfrac{T_j - T_{Li}}{T_{Mi} - T_{Li}} + \epsilon_{ij} & \text{if } T_j \le T_{Mi} \\ \psi_i \dfrac{T_j - T_{Ui}}{T_{Mi} - T_{Ui}} + \epsilon_{ij} & \text{if } T_j > T_{Mi} \end{cases}, \qquad i = 1...4, \quad j = 1...15, \qquad (8.8)$$

where $\{\epsilon_{ij}\}$ are NID$(0,\sigma^2)$ and where T_{Li}, T_{Ui} are the lower and upper temperatures at which zero growth occurs for fish in weight class i, i = 1...4. Model (8.8) is fitted by nonlinear least squares with

TABLE 8.5

Parameter estimates for models (8.7) and (8.8)

Parameter	Estimate	95% confidence limits	
ψ_1	1.298	1.248	1.344
ψ_2	0.808	0.776	0.841
ψ_3	0.676	0.628	0.724
ψ_4	0.518	0.467	0.566
T_{M1}	13.58	13.29	13.86
T_{M2}	13.42	13.15	13.75
T_{M3}	13.63	13.00	14.14
T_{M4}	13.64	12.22	14.32
b_{11}	0.1267	0.1157	0.1378
b_{12}	0.0807	0.0737	0.0875
b_{13}	0.0658	0.0545	0.0795
b_{14}	0.0521	0.0405	0.0694
b_{21}	-0.217	-0.235	-0.198
b_{22}	-0.140	-0.150	-0.132
b_{23}	-0.114	-0.133	-0.094
b_{24}	-0.087	-0.106	-0.067
T_{L1}	3.33	2.66	3.88
T_{L2}	3.41	2.93	3.81
T_{L3}	3.35	1.86	4.45
T_{L4}	3.68	1.80	4.94
T_{U1}	19.56	19.31	19.87
T_{U2}	19.19	19.01	19.37
T_{U3}	19.57	19.13	20.19
T_{U4}	19.58	18.97	20.36

rss = 0.01268 on 14 d.f.
s^2 = 0.000906

i) all the parameters unconstrained, rss = 0.01268,

ii) the constraints $T_{Mi} = T_M$, $i = 1...4$, rss = 0.01348,

iii) as ii) and the constraints $T_{Li} = T_L$, $T_{Ui} = T_U$, $i = 1...4$,

rss = 0.01981.

Hence, $T_{Mi} = T_M$, $i = 1...4$ ($F = 0.29$, $F(3,14;0.05) = 3.34$, $F_{sim}(0.05) = 3.24$, $p > 0.05$), so the temperature at which maximum growth occurs does not vary with weight. Also $T_{Li} = T_L$, $T_{Ui} = T_U$, $i = 1...4$ ($F = 1.33$, $F(6,17;0.05) = 2.70$, $F_{sim}(0.05) = 2.53$, $p > 0.05$), so the temperatures at which zero growth occurs do not vary with weight.

Let $\{W_i\}$ be the initial weights of the trout in each weight class ($W_1 = 10$, $W_2 = 50$, $W_3 = 90$, $W_4 = 250$). A plot of the least squares estimates of $\{\psi_i\}$ (obtained in iii above) against $\{W_i\}$, suggests that $\{\psi_i\}$ and $\{W_i\}$ are related by an "inverse power" relationship. Hence, model (8.8) is fitted with the constraint

iv) as iii) and the constraint $\psi_i = \xi W_i^{-\eta}$, $i = 1...4$, for parameters ξ and η,

rss = 0.02023.

Thus $\psi_i = \xi W_i^{-\eta}$, $i = 1...4$ ($F = 0.24$, $F(2,23;0.05) = 3.42$, $F_{sim}(0.05) = 2.82$, $p > 0.05$). An examination of the residuals from the final model reveals no evidence of heterogeneous variances or lack of fit. Hence, the model

$$\bar{g}_{ij} = \begin{cases} \xi W_i^{-\eta} \dfrac{T_j - T_L}{T_M - T_L} + \epsilon_{ij} & \text{if } T_j \le T_M \\ \xi W_i^{-\eta} \dfrac{T_j - T_U}{T_M - T_U} + \epsilon_{ij} & \text{if } T_j > T_M \end{cases}, \qquad i = 1...4, \quad j = 1...15, \qquad (8.9)$$

is a good description of the relationship between $\bar{g}$ and temperature and initial weight. Parameter estimates are given in Table 8.6.

The analysis above is repeated with a "global" hyperbola model

$$\bar{g}_{ij} = \psi_i\left(1+\frac{1}{2}\left((T_j-T_{Mi})\left(\frac{1}{T_{Mi}-T_{Li}}+\frac{1}{T_{Mi}-T_{Ui}}\right)-((T_j-T_{Mi})^2+c_i)^{\frac{1}{2}}\left(\frac{1}{T_{Mi}-T_{Li}}-\frac{1}{T_{Mi}-T_{Ui}}\right)\right)\right)+\epsilon_{ij},$$

$$i = 1...4, \quad j = 1...15, \qquad (8.10)$$

Model (8.10) is fitted by nonlinear least squares with the constraints i) to iv) above and in addition v) as iv) and the constraint $c_i = 0$, $i = 1...4$.

As before,

$T_{Mi} = T_M$, $i = 1...4$ ($F = 0.05$, $F(3,10;0.05) = 3.71$,

$F_{sim}(0.05) = 3.52$, $p > 0.05$),

$T_{Li} = T_L$, $T_{Ui} = T_U$, $i = 1...4$ ($F = 0.66$, $F(6,13;0.05) = 2.91$,

$F_{sim}(0.05) = 2.55$), $p > 0.05$),

$\psi_i = \xi W_i^{-\eta}$, $i = 1...4$ ($F = 2.55$, $F(2,19;0.05) = 3.52$,

$F_{sim}(0.05) = 3.35$, $p > 0.05$).

Further, $c_i = 0$, $i = 1...4$ ($F = 2.04$, $F(4,21;0.05) = 2.84$, $F_{sim}(0.05) = 2.15$, $p > 0.05$), so the hyperbola (8.10) reduces to the pair-of-straight-lines (8.9).

In the analysis so far, it has been assumed that the growth rate of an individual fish does not vary over the short time period of a sub-experiment (cf equation (8.1)), and that $\bar{g}_{ij}$ represents the average growth rate of trout of weight W_i at

TABLE 8.6

Parameter estimates for model (8.9)

Parameter	Estimate	95% confidence limits	
ξ	2.58	2.37	2.75
η	0.294	0.274	0.315
T_M	13.67	13.49	13.84
T_L	3.31	2.99	3.61
T_U	19.35	19.23	19.49

rss = 0.0202 on 25 d.f.
s^2 = 0.000808

temperature T_j. However, in some sub-experiments, the increase in weight of the trout was quite large. For example, the 50-g trout kept at 12.8°C each weighed over 70 g by the end of the sub-experiment. Since growth rates decrease as fish weight increases, the growth rates of these trout will have decreased slightly during the sub-experiment (although not sufficiently for this to have been detected when fitting model (8.1)). In the next Section, the growth rate of a fish is allowed to vary continuously as the fish grows. However, as an intermediate step, it is interesting to note the effect of using the average weight of the trout during a sub-experiment as a covariate in the analysis, rather than the initial weight of the trout.

Let

i) w_{ij1} be the average weight of the trout in weight class i at the start of the sub-experiment at temperature T_j, i = 1...4, j = 1...15,

ii) w_{ij2} be the average weight of the trout in weight class i at the end of the sub-experiment at temperature T_j, i = 1...4, j = 1...15.

Then

$$\bar{w}_{ij} = (w_{ij1} + w_{ij2})/2$$

is approximately the average weight of the trout in weight class i during the sub-experiment at temperature T_j, i = 1...4, j = 1...15. The model

$$\bar{g}_{ij} = \begin{cases} \xi(\bar{w}_{ij})^{-\eta}\dfrac{T_j - T_L}{T_M - T_L} + \epsilon_{ij} & \text{if } T_j \le T_M \\ \xi(\bar{w}_{ij})^{-\eta}\dfrac{T_j - T_U}{T_M - T_U} + \epsilon_{ij} & \text{if } T_j > T_M \end{cases}, \quad i = 1...4, \quad j = 1...15, \qquad (8.11)$$

where $\{\varepsilon_{ij}\}$ are NID$(0,\sigma^2)$, is an excellent description of the data and, in particular, has a smaller residual sum of squares than model (8.9) (model (8.11): rss = 0.0186, model (8.9): rss = 0.0202). Further, model (8.11) describes the data for 50-g trout almost as well as the hyperbola (8.6) (model (8.11): rss = 0.00852, model (8.6): rss = 0.00835), (Fig. 8.7). Thus, it appears that slight changes in the growth rates of the 50-g trout at temperatures close to T_M are responsible for spuriously indicating that the hyperbola is a better description of the data for 50-g trout than the pair-of-straight-lines. For this reason, and because the global hyperbola model is not a significantly better fit than the global pair-of-straight-lines model, the hyperbola is not considered further.

8.5 A GENERAL MODEL OF GROWTH

Assume, for one moment, that models (8.1) and (8.9) are deterministic. Over short time periods, growth is well described by (8.1). That is, the weights $w(t)$ and $w(t + \delta t)$ of a trout at times t and $t + \delta t$ are related by

$$\log(w(t+\delta t)) = \log(w(t)) + \frac{g(w(t),T)}{100}\delta t + O(\delta t^2),$$

where

i) $g(w(t),T)$ is the growth rate of a trout of weight $w(t)$ at temperature T,

ii) the term $O(\delta t^2)$ incorporates the effect of slight changes in the growth rate between t and $t + \delta t$.

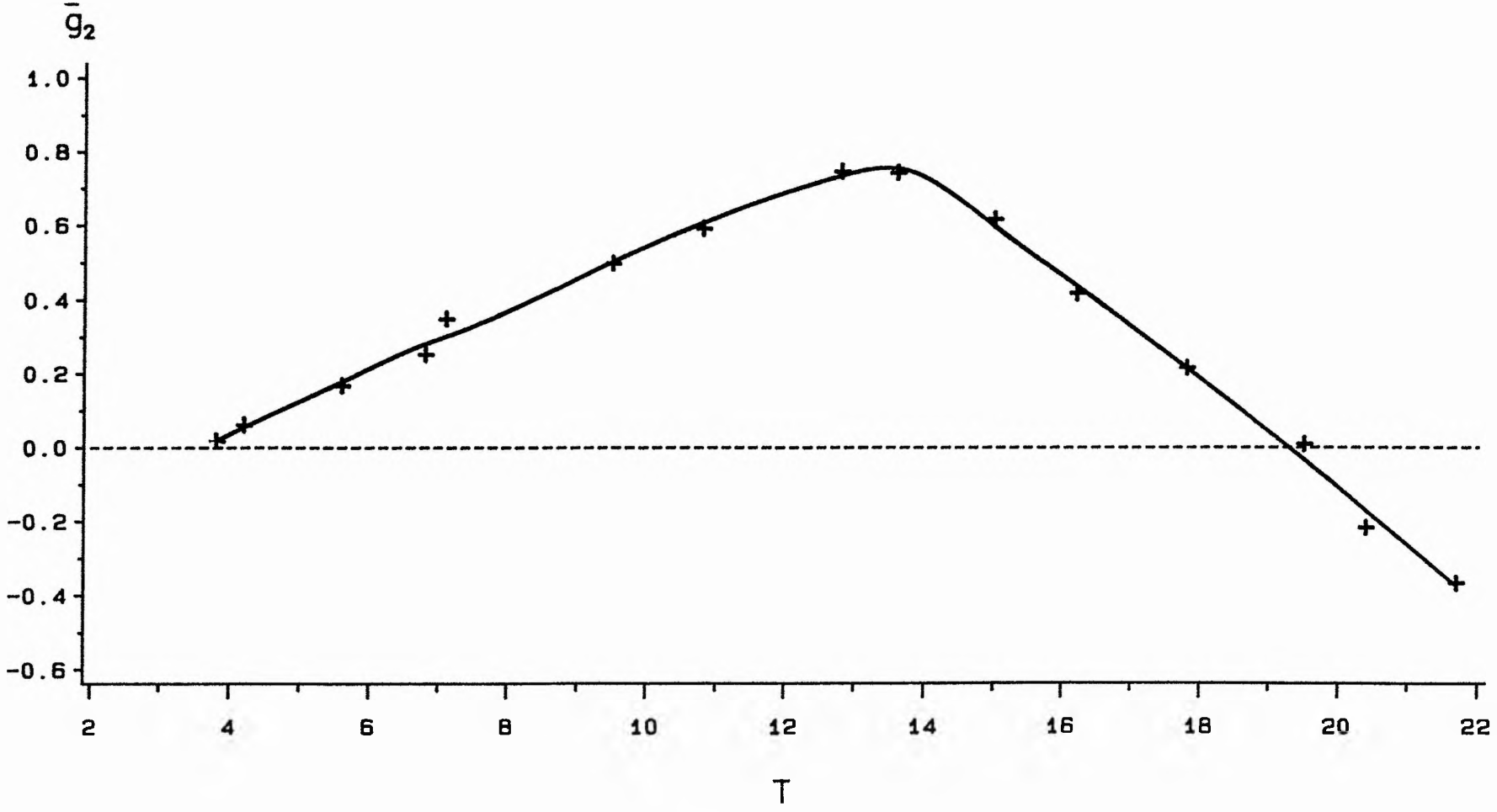


Fig. 8.7 The pair–of–straight–lines model with a weight covariate (model (8.11)) fitted to the 50–g trout data.

In the limit as $\delta t \to 0$,

$$\frac{dw(t)}{dt} = w(t)\frac{g(w(t),T)}{100}. \tag{8.12}$$

Inserting (8.9) into (8.12) gives

$$\begin{aligned}\frac{dw(t)}{dt} &= w(t)w(t)^{-\eta}h(T) \\ &= w(t)^{1-\eta}h(T),\end{aligned}$$

where

$$h(T) = \begin{cases} \dfrac{\xi(T-T_L)}{100(T_M-T_L)} & \text{if } T \le T_M \\ \dfrac{\xi(T-T_U)}{100(T_M-T_U)} & \text{if } T > T_M \end{cases}. \tag{8.13}$$

This integrates to give

$$w(t) = \{w(0)^{\eta} + \eta h(T)t\}^{1/\eta} \tag{8.14}$$

Equation (8.14) is a (deterministic) growth model which relates the weight of a trout at time t to the weight of the trout at some previous time 0. Further, (8.14) is constructed (by integrating (8.13)) in such a way that the growth rate of a trout is allowed to vary continuously as the trout grows.

A stochastic component can be introduced into (8.14) in a number of ways. A first approach is to consider an additive error with constant variance. The model

$$w_{ij2} = \{(w_{ij1})^{\eta} + \eta h(T_j)t_j\}^{1/\eta} + \epsilon_{ij}, \qquad i = 1...4, \quad j = 1...15, \qquad (8.14a)$$

where $\{\epsilon_{ij}\}$ are NID$(0, \sigma^2)$, is fitted by nonlinear least squares regression and is an excellent description of the data. An examination of the residuals gives no evidence that the form of the error term is inappropriate. However, it is arguable that a multiplicative error term might be more realistic biologically, since the weight of a large trout changes more than that of a small trout during the same sub-experiment; for example, in the experiment at 12.8°C, the 10-g trout increased their weight by about 10g, whereas the 250-g trout increased their weight by about 50 g. The model

$$\log(w_{ij2}) = \frac{1}{\eta}\log\{(w_{ij1})^{\eta} + \eta h(T_j)t_j\} + \epsilon_{ij}, \quad i = 1...4, \quad j = 1...15, \qquad (8.14b)$$

where $\{\epsilon_{ij}\}$ are NID$(0, \sigma^2)$, again fitted by nonlinear least squares regression, is also an excellent description of the data and the residuals reveal no pattern of heterogeneous variances. There is little to choose between the two models. In particular, although the parameter estimates from the two models are slightly different, the fitted values are very similar. Hence, it is concluded that both (8.14a) and (8.14b) are good models of the growth in weight of brown trout fed on maximum rations. Parameter estimates are given in Table 8.7.

In model (8.14a), the least squares estimate of T_M is relatively close to the design point at 13.6°C (since a linear approximation 95% confidence interval for T_M includes 13.6). Hence, the nonlinear behaviour of the model is likely to be large. This is

TABLE 8.7

Parameter estimates for model (8.14)

Model (8.14a)

Parameter	Estimate	Simulation 95% confidence limits	
ξ	2.628	2.332	2.925
η	0.290	0.268	0.311
T_M	13.53	13.41	13.67
T_L	3.62	3.39	3.85
T_U	19.51	19.39	19.64

rss = 12.88 on 25 d.f.
s^2 = 0.515

Model (8.14b)

Parameter	Estimate	Lin approx					
		s.e.	correlations				
ξ	3.098	0.131	1.00				
η	0.327	0.011	0.97	1.00			
T_M	13.41	0.083	0.07	0.07	1.00		
T_L	3.54	0.124	0.09	0.00	-0.26	1.00	
T_U	19.41	0.067	-0.02	-0.00	-0.07	0.00	1.00

rss = 0.00249 on 25 d.f.
s^2 = 0.0000996

confirmed by a simulation, which shows that the distribution of $\hat{T}_M$ is multimodal. Therefore, only simulation confidence intervals are given in Table 8.7. However, in the case of model (8.14b), a linear approximation 95% confidence interval for T_M does not include 13.6. A simulation confirms that the distribution of $\hat{T}_M$ is effectively confined to the interval (12.8, 13.6), so the change in solution locus segment at $T_M = 13.6$ is likely to have little effect on the nonlinear behaviour of the model. Hence, the standard measures of nonlinearity can be applied (Table 8.8). The intrinsic nonlinearity of the model is small. However, the maximum parameter effects curvature is slightly greater than the critical value of $1/(2\sqrt{F})$ (although the RMS parameter effects curvature is considerably lower than $1/(2\sqrt{F})$). A variety of parameter transformations was investigated, but none reduced the parameter effects curvature. In particular, the reparameterisation

$$\log(w_{ij2}) = \frac{1}{\eta}\log\{(w_{ij1})^{\eta} + h'(T_j)t_j\} + \epsilon_{ij}, \qquad i = 1\ldots4, \quad j = 1\ldots15,$$

where

$$h'(T) = \begin{cases} \dfrac{\xi'(T - T_L)}{100(T_M - T_L)} & \text{if } T \le T_M \\ \dfrac{\xi'(T - T_U)}{100(T_M - T_U)} & \text{if } T > T_M \end{cases},$$

and

$$\xi' = \eta\xi,$$

TABLE 8.8

Measures of nonlinearity for model (8.14b)

	max	RMS
Intrinsic curvature	0.03	0.01
Parameter effects curvature	0.38	0.11

$$1/(2\sqrt{F(5,25;0.05)}) = 0.31$$

Parameter	ξ	η
Least squares estimate Lin approx s.e.	3.098 0.131	$3.274*10^{-1}$ $0.112*10^{-1}$
Box's bias	0.003	$4*10^{-4}$
Hougaard's skewness Lowry and Morton's λ	0.128 0.001	0.021 $4*10^{-4}$
Simulation measures: True value Variance of l.s.e. (lin approx)	 3.098 $1.716*10^{-2}$	 $3.274*10^{-1}$ $1.254*10^{-4}$
Simulated value Simulated variance Skewness stat Excess kurtosis stat	3.095 $1.903*10^{-2}$ 0.66 -0.93	$3.270*10^{-1}$ $1.362*10^{-4}$ 0.92 -0.83
Lowry and Morton's λ	0.001	$4*10^{-4}$

TABLE 8.8 (continued)

Parameter	T_M	T_L
Least squares estimate	13.406	3.544
Lin approx s.e.	0.083	0.124
Box's bias	-0.0003	-0.002
Hougaard's skewness	-0.026	-0.113
Lowry and Morton's λ	$3*10^{-4}$	0.001
Simulation:		
True value	13.406	3.544
Variance of l.s.e. (lin approx)	$6.922*10^{-3}$	$1.538*10^{-2}$
Simulated value	13.411 *	3.450
Simulated variance	$6.259*10^{-3}$ *	$1.439*10^{-2}$
Skewness stat	-1.83	0.49
Excess kurtosis stat	0.94	0.33
Lowry and Morton's λ	0.004	0.001

Parameter	T_U
Least squares estimate	19.406
Lin approx. s.e.	0.067
Box's bias	-0.001
Hougaard's skewness	-0.083
Lowry and Morton's λ	0.001
Simulation:	
True value	19.406
Variance of l.s.e. (lin approx)	$4.476*10^{-3}$
Simulated value	19.409
Simulated variance	$4.447*10^{-3}$
Skewness stat	0.92
Excess kurtosis stat	1.13
Lowry and Morton's λ	0.001

* indicates moderate nonlinear behaviour
** indicates serious nonlinear behaviour

has much greater parameter effects curvature than model (8.14b). Since, the simulation measures indicate that the nonlinear behaviour of (8.14b) is small, it is concluded that (ξ, η, T_M, T_L, T_U) is an adequate parameterisation of (8.14b) and linear approximation standard errors and correlations are given in Table 8.7.

8.6 DISCUSSION

The growth model (8.14) is an excellent description of the growth in weight of brown trout fed on maximum rations. The model can be applied to trout over a wide range of weights (10 - 300 g) and a wide range of water temperatures (3.8 - 21.7°C). Further, the model has only a small number of parameters, of which three, T_M, T_L and T_U, have the desirable biological interpretations of the temperatures at which maximum and zero growth occurs.

The one unsatisfactory feature of the growth model is the presence of a change point, since biological arguments suggest that the sudden change in growth patterns in the region of maximum growth is very implausible. More realistic models could be formulated, either empirically, or on the basis of theoretical arguments. However, to achieve both an adequate description of the data and biological realism, such models would have to be more complex than (8.14); for example, a growth model based on the hyperbola would have more parameters than (8.14). Since model (8.14) is perfectly adequate for the observed data, fitting more complex models is likely to lead to near collinearity (in the matrix of first partial derivatives) and hence unstable parameter estimates and forecasts (Simonoff and Tsai, 1989). Hence, a change-point model appears to be necessary if the observed data are to be described in a parsimonious way.

CHAPTER 9 A LENGTH CLASS MODEL OF SURVIVAL FROM t_1 TO t_2

9.1 INTRODUCTION

The relationships between the number of 0+ parr R_1 at time t_1 (the May/June electro-fishing) and the number of eggs E and between the number of 0+ parr R_2 at time t_2 (the August/September electro-fishing) and E are well described by models (6.6) and (6.7) respectively. These models show that the survival rate between t_1 and t_2 still depends on the initial number of eggs and that R_2 can not be predicted from R_1 alone. However, it is unlikely that E has a direct effect on the survival rate. A more plausible explanation is as follows. The probability that an individual trout survives between t_1 and t_2 will vary between trout. In addition to the density of trout at t_1 (ie R_1), the survival probabilities could depend on a number of covariates, such as the state of health of the trout at t_1, the size of the trout at t_1, etc. The survival rate between t_1 and t_2 will then appear to depend on E if one of these covariates is egg dependent.

The length distribution of the trout population at t_1 is egg dependent. Further, survival probabilities might be expected to depend on length since

i) large trout might compete for food and territories more effectively than small trout,

ii) length might be a good indicator of the state of health of a trout. Hence, this Chapter describes a way of modelling the survival of an individual trout between t_1 and t_2 in terms of its length at t_1.

The development of a model that relates R_2 to the population at t_1 has two other advantages. First, by introducing many more states that describe the "health" of the population at t_1, the life stages during the first summer of life can be regarded as Markovian; survival between t_0 and t_1 does not depend directly on anything that occurs before t_0 and survival between t_1 and t_2 does not depend directly on anything that occurs before t_1. Secondly, the model could be useful in the study of other fish populations, particularly marine species, for which it is easier to sample the number of juvenile fish than the number of eggs.

The problems of relating R_2 to R_1 are described more fully in Section 9.2. The length class model is developed in Section 9.3 and the model is applied to the Black Brows Beck population in Section 9.4. In Section 9.5, a set of simulations is described that investigates some of the properties of the model.

9.2 THE RELATIONSHIPS BETWEEN E, R_1 AND R_2

9.2.1 Ricker Models

The relationship between R_1 and E is well described by the Ricker model

$$R_{1i} = E_i \exp(\alpha_1 - \beta_1 E_i) + \epsilon_i, \qquad i = 67 \ldots 84, \tag{9.1}$$

where $\{\epsilon_i\}$ are NID$(0,\sigma_1^2)$ (Section 4.1). Similarly, the relationship between R_2 and E (with the exception of the 1984 data point) is well described by a Ricker model modified to incorporate the effect of summer drought

$$R_{2i} = E_i \exp(\alpha_2 - \beta_2 E_i - \zeta_2/(w_i - \xi_2)) + \epsilon_i, \qquad i = 67 \ldots 83, \tag{9.2}$$

where (ε_i) are NID$(0,\sigma_2^2)$ (Section 6.2.5).

The number of eggs E can be eliminated from models (9.1) and (9.2) as follows. For simplicity, it is assumed that the models are both deterministic and that

$$g(w_i) = \exp(-\xi_2/(w_i - \xi_2)), \qquad i = 67...83.$$

If $\beta_1 = \beta_2$, R_2 can be written uniquely in terms of R_1 as

$$R_{2i} = g(w_i)\exp(\alpha_2 - \alpha_1)R_{1i}, \qquad i = 67...83. \tag{9.3}$$

In this case, there is proportional survival between t_1 and t_2 with a constant of proportionality

$$g(w_i)\exp(\alpha_2 - \alpha_1), \qquad i = 67...83.$$

If $\beta_1 \neq \beta_2$, then

$$R_{2i}/R_{1i} = g(w_i)\exp(\alpha_2 - \alpha_1 - (\beta_2 - \beta_1)E_i), \qquad i = 67...83,$$

so that

$$E_i = \frac{1}{(\beta_2 - \beta_1)}\{\log(g(w_i)R_{1i}/R_{2i}) + \alpha_2 - \alpha_1\}, \qquad i = 67...83. \tag{9.4}$$

Inserting (9.4) in (9.2) and rearranging the equations gives

$$\frac{1}{(\beta_2-\beta_1)}\{\log(g(w_i)R_{1i}/R_{2i})+\alpha_2-\alpha_1\}\{g(w_i)^{-\beta_1}R_{1i}^{-\beta_2}R_{2i}^{\beta_1}\}^{1/(\beta_2-\beta_1)}\exp\left(\frac{\alpha_1\beta_2-\alpha_2\beta_1}{\beta_2-\beta_1}\right)=1 \qquad (9.5)$$

In this case, R_2 can not be expressed uniquely as a function of R_1.

9.2.2 Horseshoe Curves

In this section the case $\beta_1 \neq \beta_2$ is examined further. Suppose $g(w_i) = 1$, so that there is no drought effect. Equation (9.5) simplifies to

$$\frac{1}{(\beta_2-\beta_1)}\{\log(R_{1i}/R_{2i})+\alpha_2-\alpha_1\}\{R_{1i}^{-\beta_2}R_{2i}^{\beta_1}\}^{1/(\beta_2-\beta_1)}\exp\left(\frac{\alpha_1\beta_2-\alpha_2\beta_1}{\beta_2-\beta_1}\right)=1, \quad i=67...83.$$

R_2 can now be drawn as a function of R_1 to give a "horseshoe" curve (Fig. 9.1). The form of this curve can be explained by examining a graph of the original Ricker models (9.1) and (9.2), again assuming $g(w_i) = 1$ (Fig. 9.2). A value of R_1, R_1^* say, can come from two possible values of E, E^* and E^{**} say, and since $\beta_1 \neq \beta_2$, these numbers of eggs produce different numbers R_2^* and R_2^{**} of 0+ parr at t_2. Thus each value of R_1 is associated with two values of R_2.

If $\beta_1 > \beta_2$, as for the Black Brows Beck population, the "upper branch" of the horseshoe corresponds to larger values of E than the "lower branch". In this case, there is an inconsistency in the original Ricker models, since for very high values of E, there would be greater than 100% survival between t_1 and t_2 (ie $R_2 > R_1$). However, the parameter estimates for Black Brows Beck indicate that this will only occur if $E > 12{,}200$ - a much higher value than any recorded - so that the problem

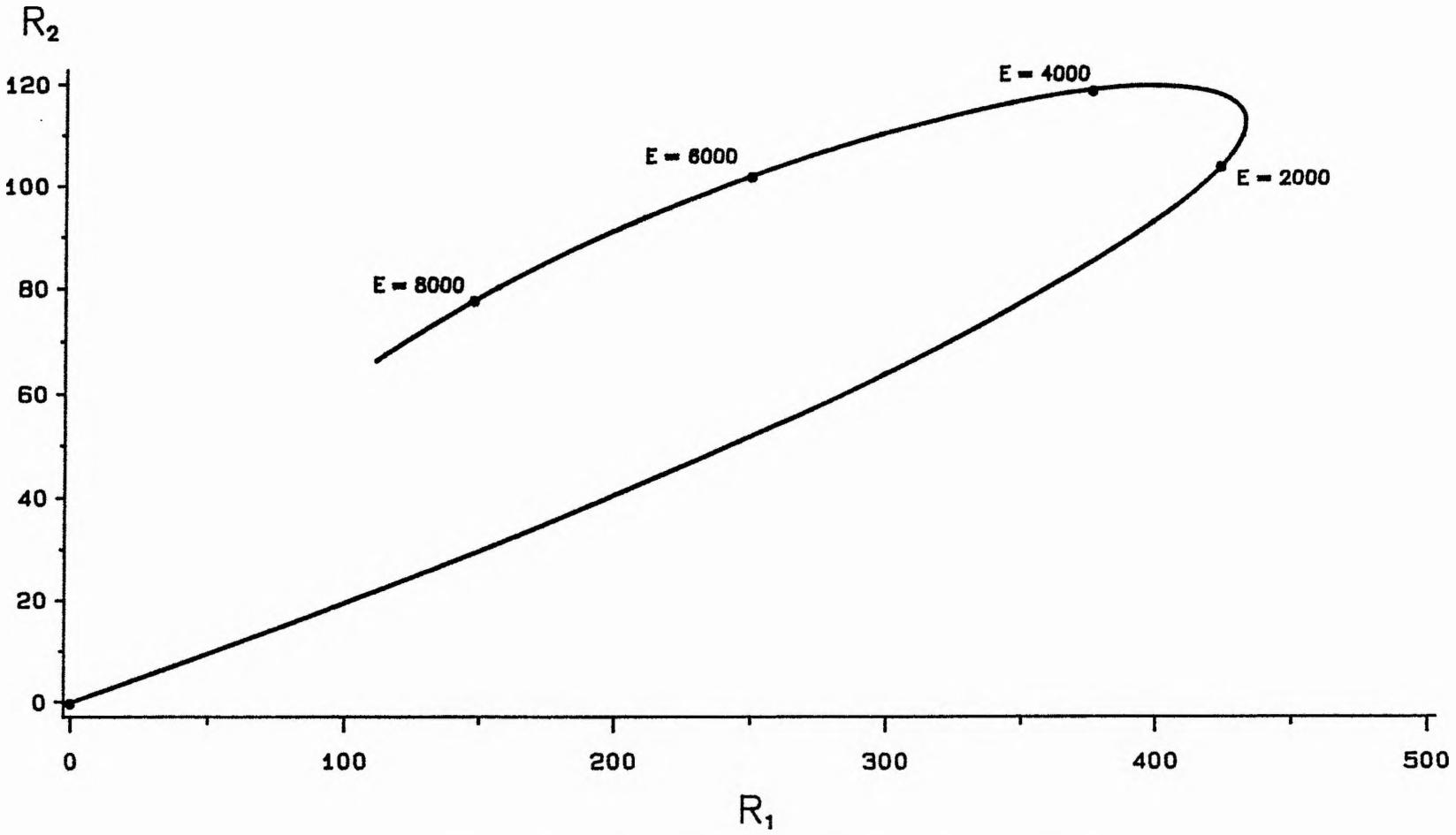


Fig. 9.1 The 'horseshoe' relationship between R_2 and R_1

($\alpha_1 = -0.74$, $\alpha_2 = -2.35$, $\beta_1 = 4.1 * 10^{-4}$, $\beta_2 = 2.8 * 10^{-4}$)

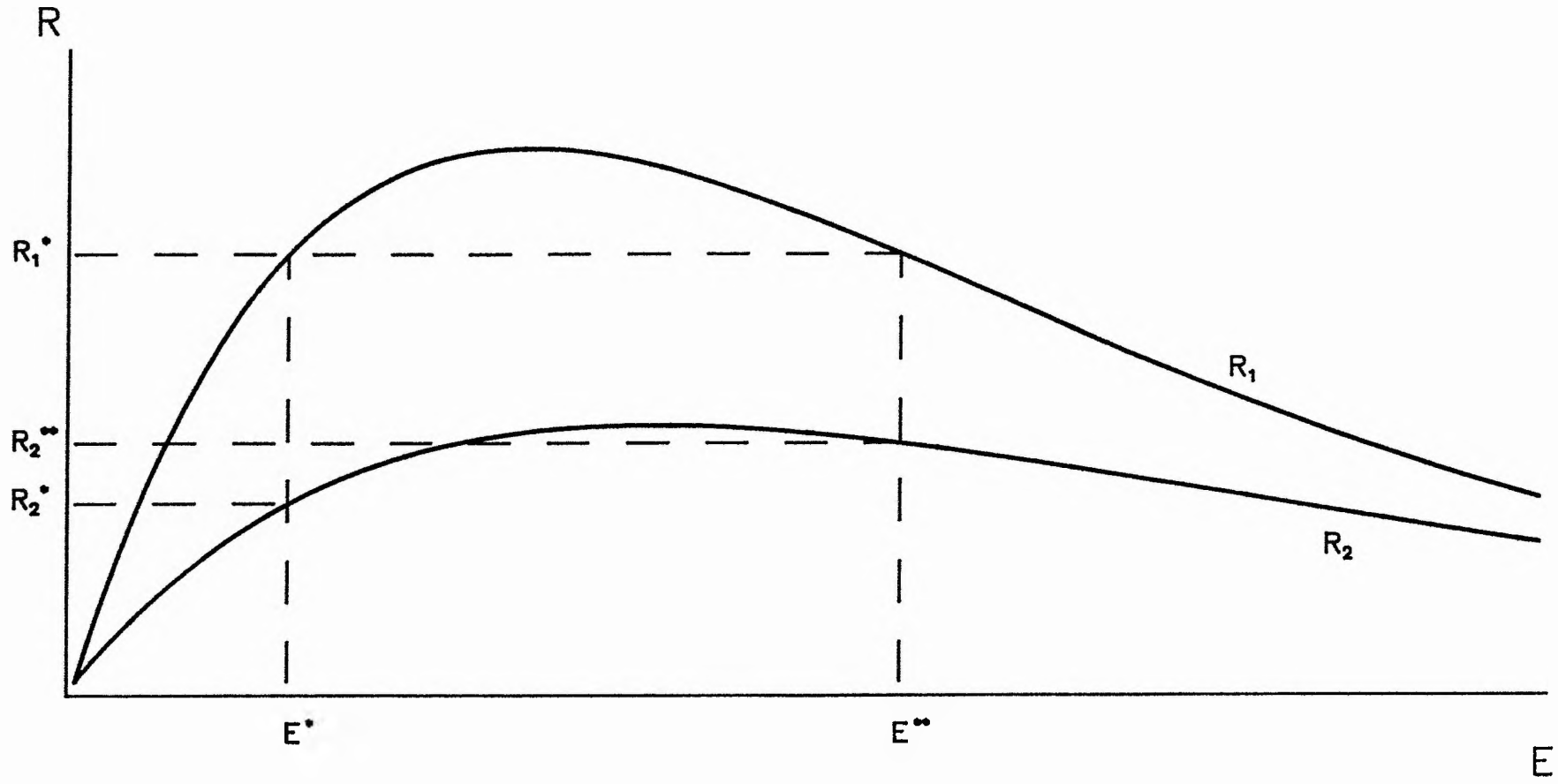


Fig. 9.2 Ricker curves describing the relationship between R_1 and E and between R_2 and E when $\beta_1 \neq \beta_2$.

can effectively be ignored. If $\beta_1 < \beta_2$, the upper branch of the horseshoe corresponds to smaller values of E than the lower branch and there are no inconsistencies in the original Ricker models.

As $(\beta_1 - \beta_2) \rightarrow 0$, the two branches of the horseshoe close together. In the limit, they merge to give a straight line, giving proportional survival between t_1 and t_2; this is consistent with equation (9.3).

9.2.3 May/June Length Distributions

Since $\beta_1 \neq \beta_2$ for the Black Brows Beck population (Section 6.2.7), it is not possible to relate R_2 to R_1 alone. More information is needed to model the survival rate between t_1 and t_2 in terms of the population at t_1.

One possible approach would be to take an egg dependent attribute of the population at t_1 which, for a given value of R_1, determines which branch of the horseshoe is used to estimate the corresponding value of R_2. The only suitable available data is the length distribution at t_1. The length of each trout caught was measured during the May/June electro-fishing. Although the mean length is not correlated with E, the range of the lengths is negatively correlated with E (Elliott, 1984b); it is probable that as E increases, the additional competition for food reduces both the survival rate of smaller trout and the growth rate of larger trout between t_1 and t_2. This method was investigated briefly, but did not prove to be successful.

The approach developed in this Chapter is to consider the trout population one year at a time and to investigate whether the probability that a trout survives between t_1 and t_2 depends on its length at t_1. A model is developed which describes, for

each year separately, the "transition" of the population between t_1 and t_2 and which incorporates length dependent survival probabilities. The survival probabilities are estimated and the variation of the probabilities over the years is investigated.

It is important to emphasise that none of the trout were marked, since marking is liable to alter survival rates. Therefore, there is no immediate way of estimating the survival probability of a trout of a given length.

9.3 THE LENGTH CLASS MODEL

9.3.1 Multimodal Length Distributions

The length distributions at t_1 and t_2 show no common pattern over the years, and in general, are multimodal. For example, a histogram of the lengths at t_1 in 1980 contains about five peaks (Fig. 9.3). It is not possible to account for the multimodality using any of the available data. In particular, there is no obvious relationship between the number of peaks and the number of redds; for example, there was only one redd in the study section in 1980. Thus, the length distributions at t_1 and t_2 can not be described by a parametric model with a small number of parameters. Instead, the frequency distributions of the lengths at each time period are used as unmodelled data.

All the parameters, structures and data described below are year dependent. However, for simplicity, one particular year is considered and all suffices indicating the year are omitted.

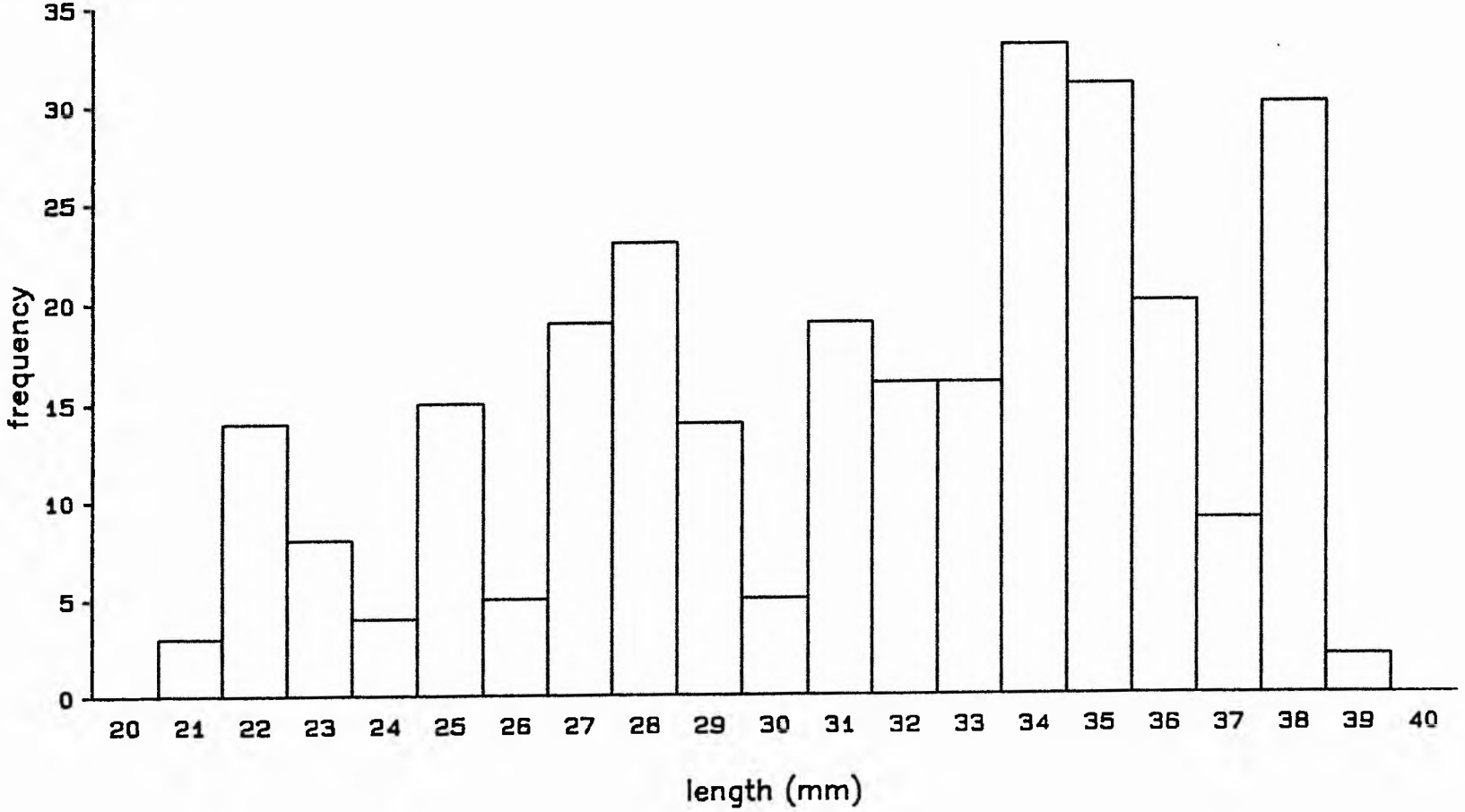


Fig. 9.3 The length distribution of the 0+ parr at t_1 in 1980.

9.3.2 May/June Length Classes

The lengths of the trout at t_1 were measured to the nearest millimetre (Elliott, 1984b). Thus, it is convenient to partition the population at t_1 into 60 May/June length classes K_i, i = 1...60, where a trout of length i mm is assigned to class K_i. That is, assuming no measurement error, K_i contains all the trout with a measured length of between (i - 0.5) mm and (i + 0.5) mm. Since all the 0+ parr at t_1 have a measured length of between 21 mm and 40 mm, the classes K_i, i = 1...20,41...60, are always empty; however, they are included for notational convenience.

9.3.3 Summer Growth

It is necessary to partition the population at t_2 into August/September length classes that correspond in some way to the May/June length classes $\{K_i\}$. One obvious partition would be to split the population into 1 mm size groups. However, a disadvantage of this approach is that the growth of the trout between t_1 and t_2 varies considerably between years due to the effect of varying temperature conditions. For example, the lengths of the trout at t_2 in 1970 ranged from 38 mm to 75 mm, whereas in 1978, the lengths ranged from 48 mm to 84 mm. Therefore, the August/September length classes would not correspond in any natural way to the May/June length classes; in 1970, fish would transfer to the "smaller" August/September length classes whereas in 1978, fish would transfer to the "larger" August/September length classes. An alternative approach, described below, uses the growth model developed in Chapter 8 to eliminate the effect of variation in growth due to different water temperatures. The August/September length classes are defined so that, other things being equal, trout from a particular May/June length class will transfer to the same August/September length class regardless of the water temperature between t_1 and t_2.

Suppose a trout has length l_1 at t_1. The weight of the trout w_1 is estimated from the equation

$$\log(w_1) = a_1 + b_1 \log(l_1), \tag{9.6}$$

where the values of the parameters a_1 and b_1 are estimated by ordinary least squares regression from the lengths and weights of a small number of trout taken at t_1 from the lower study section of the stream (Elliott, 1984b).

The model

$$\log(w(t)) = \frac{1}{\eta} \log\{w(0)^{\eta} + \eta h(T) t\} + \epsilon, \tag{9.7}$$

where

$$h(T) = \begin{cases} \dfrac{\xi(T - T_L)}{100(T_M - T_L)} & \text{if } T \leq T_M \\ \dfrac{\xi(T - T_U)}{100(T_M - T_U)} & \text{if } T > T_M \end{cases}$$

describes the growth of brown trout feeding on maximum rations over a time period t at water temperature T (Section 8.5). Equation (9.7) is derived from experimental data on larger trout than the 0+ parr in Black Brows Beck. However, it is assumed that it can be applied to the Black Brows Beck population and can be used to estimate the weight of a trout at t_2 that has fed on maximum rations between t_1 and t_2. Since approximate water temperatures were measured at half-monthly intervals, the period

between t_1 and t_2 is divided into small time intervals and the weight of a trout at the end of each interval is estimated iteratively from equation (9.7) and the weight at the start of the interval. For example, suppose that t_1 is 23 May and that the water temperature over the period 16 to 31 May is T_1. The weight on 1st June of the trout of initial weight w_1 is estimated from

$$\log(w(1\ \text{June})) = \frac{1}{\eta}\log\{w_1^{\eta} + (32-23)\eta h(T_1)\}.$$

The weight on 16 June, 1 July etc is then estimated until the weight w_2 at t_2 is found.

Finally, the length of the trout l_2 at time t_2 is estimated from the equation

$$\log(l_2) = a_2 + b_2\log(w_2), \tag{9.8}$$

where a_2 and b_2 are estimated by ordinary least squares regression from samples taken at t_2 from the lower study section of the stream. (The correlations between log-length and log-weight in (9.6) and (9.8) are so good that the calibration problem can be ignored).

In this way, a growth function $f(.)$ is constructed each year that maps the length of a fish at t_1 to its corresponding theoretical length at t_2 assuming that the growth in weight of the fish is given by (9.7).

9.3.4 August/September Length Classes

The population at t_2 is partitioned into 60 August/September length classes L_j, j = 1...60, where a trout of length between $f(j - 0.5)$ mm and $f(j + 0.5)$ mm at t_2 is assigned to L_j. The length classes in August/September correspond naturally to those in May/June as follows. If a trout is in K_i at t_1, it has a length of between $(i - 0.5)$ mm and $(i + 0.5)$ mm; if the trout survives the summer, it will have a theoretical length of between $f(i - 0.5)$ mm and $f(i + 0.5)$ mm and will be in L_i at t_2 (Fig. 9.4). Fish that die between t_1 and t_2 are assigned to a special length class L_*.

The lengths of the trout caught at t_2 are measured to the nearest mm. However, in general, the range of L_j is not of the form $(k - 0.5)$ mm to $(k + 0.5)$ mm for some integer k, so a trout could be assigned to one of a number of length classes. For example, if L_j contains trout of true length between 62.4 mm and 64.1 mm and L_{j+1} contains trout of true length between 64.1 mm and 65.5 mm, then a trout of measured length 64 mm could be assigned to either L_j or L_{j+1}. Hence, the partitioning of trout into August/September length classes is randomised. Suppose L_j contains all trout of true lengths between Υ_1 mm and Υ_2 mm. A trout of measured length k mm is assigned to L_j with a probability equal to the proportion of the interval $(k - 0.5, k + 0.5]$ that overlaps with $(\Upsilon_1, \Upsilon_2]$; formally this probability is

$$\max\{0, [\min((k+0.5), \gamma_2) - \max((k-0.5), \gamma_1)]\}.$$

In the numerical example above, a trout of measured length 64 mm is allocated to L_j with probability 0.6 and to L_{j+1} with probability 0.4.

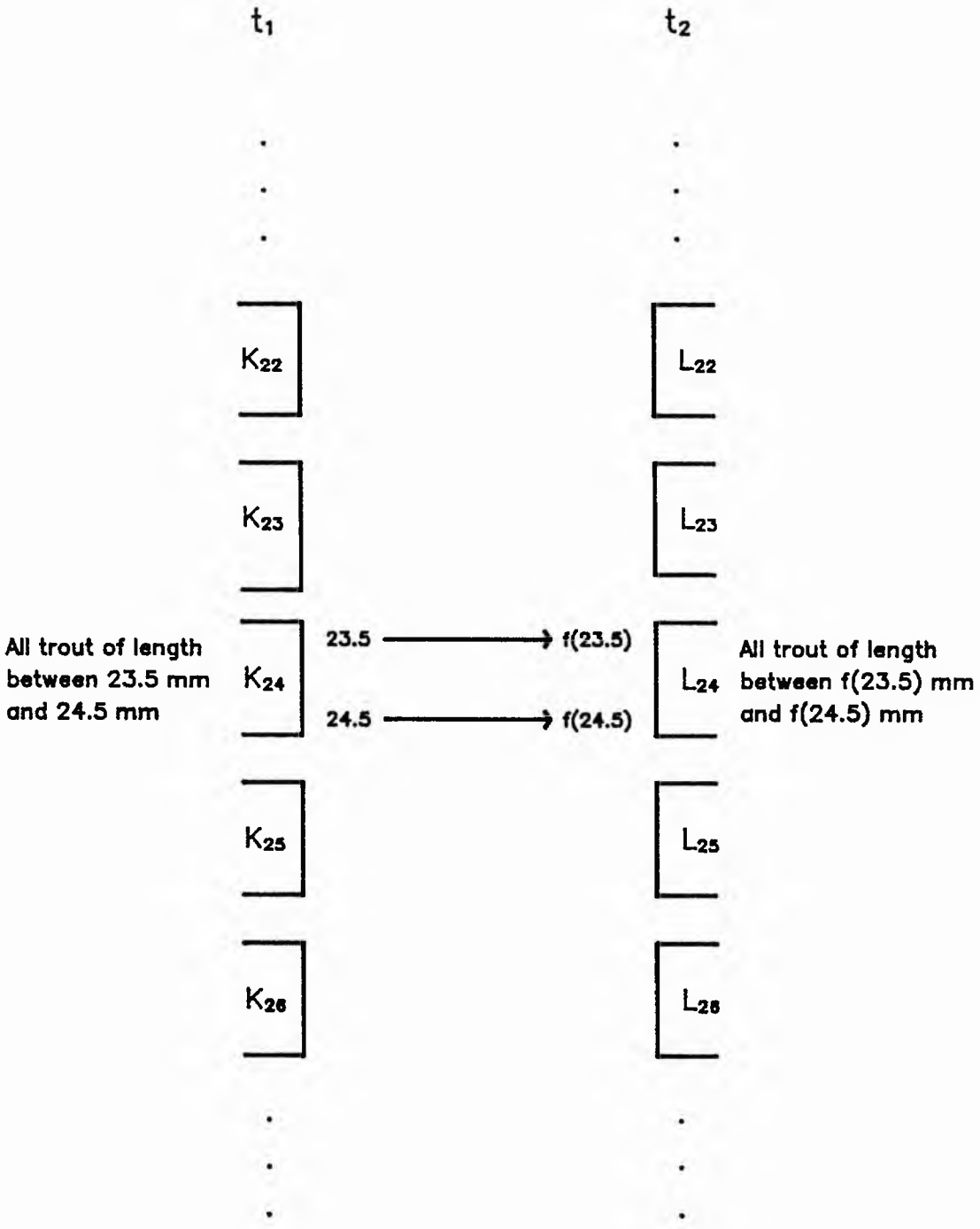


Fig. 9.4 Aug/Sep length class construction.

Let x_i be the number of trout in K_i and y_i the number of trout in L_i, $i = 1...60$. The number of trout that died between t_1 and t_2 is not known but is estimated by

$$y_\cdot = \sum_{i=1}^{60} x_i - \sum_{i=1}^{60} y_i = R_1 - R_2$$

Table 9.1 shows the allocation of trout to the May/June and August/September length classes for the years 1967 to 1983.

The advantage of the August/September length class construction above is that it incorporates the effect of varying temperature conditions. Also, it provides a qualitative method of assessing the applicability of the growth model (9.7) to the Black Brows Beck population. A comparison each year of the August/September length class containing the largest trout at t_2 with the May/June length class containing the largest trout at t_1 indicates if trout are growing larger than predicted. Should this happen frequently, the growth of the trout in Black Brows Beck can not follow the maximum growth model (9.7). Trout grow much larger than expected only in 1977 and 1979; no explanation could be found for this. However, in nearly all cases, the largest trout at t_2 is slightly smaller than predicted; this is consistent with the trout feeding on ration sizes below maximum and there is no evidence that model (9.7) can not be used. In 1983, the largest trout at t_2 is much smaller than predicted, showing the serious effect of the drought in that year.

9.3.5 Trimming the Tails

The analysis that follows involves the minimisation of a goodness of fit measure that compares the expected and observed number of trout in each August/September length class. Since the tails of the length distributions at t_2 could unduly influence

TABLE 9.1

The allocation of trout to May/June and Aug/Sep length classes

	1967		1968		1969		1970		1971		1972		1973		1974		1975		1976		1977		1978		1979		1980		1981		1982		1983	
i	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L
7				2																														
8				1																														
9				1				2																										
10				1				1						4																				
11				0		1		1						1												2								
12		1		0		1		5						1									1			4								
13		2		0		1		2						2					4				2			2		1						
14		0		0		0		3						2					5				4			1		1					1	
15		1		0		1		2				3		1			2		6			2	1			1		1					1	
16		3		1		2		1				1		0			0		2			1	2			2		1					1	
17		6		3		0		2		2		1		2			0		2			3	1			0		2					2	
18		4		1		0		4		4		3		8			3		4			1	0			0		1					1	
19		4		4		0		2		3		6		5		5	0		1			0	5			0		4					0	
20		3		1		6		2		4		3		8		0	7		3			0	4			1		1		2			5	
21	2	7	2	7	3	2		3		9	4	6	4	4		3	0		7			0	5			3	2	8		4	5		1	
22	1	12	10	6	9	6		5		3	9	11	6	3		1	0		2			1	3	2	3	14	0	2		5	10		4	
23	7	5	8	1	11	2	1	4	4	9	26	5	23	4		2	5	7	4			9	9	2	27	7	8	1	1	2	9		7	
24	2	8	4	2	7	2	4	3	1	10	10	5	6	3		0	5	8	11			1	12	1	23	1	4	3	6	2	15		3	
25	7	7	12	9	4	4	31	2	6	3	15	3	4	3		9	7	9	19	3	1	2	10	2	18	4	15	2	6	11	6		2	
26	21	5	29	1	12	5	23	3	12	4	38	8	11	9	1	5	5	13	12	8	3	5	2	2	9	3	5	6	9	5	49		3	
27	6	5	6	8	10	7	13	4	17	9	51	3	63	15	2	8	9	8	24	3	4	2	11	6	23	2	19	2	11	5	34		1	
28	75	4	36	6	33	8	12	4	56	7	39	9	24	6	9	8	39	2	45	3	7	3	35	4	8	1	23	2	24	7	51		1	
29	36	3	35	6	23	4	39	5	48	10	63	6	52	2	23	8	37	7	47	3	34	9	61	4	12	3	14	2	18	3	43		1	
30	18	5	11	3	20	3	48	6	17	8	9	6	36	3	8	12	11	6	18	2	15	4	13	5	8	4	5	4	12	6	39		0	
31	45	0	24	2	15	3	12	5	55	8	41	4	55	11	39	12	24	7	56	4	54	3	11	8	10	2	19	0	42	13	9	3	19	5
32	27	16	58	8	45	5	23	9	43	3	23	4	29	4	17	1	29	9	17	11	36	11	35	2	21	3	16	0	27	2	57	5	33	4
33	22	11	35	7	66	5	37	2	26	7	68	3	13	4	19		14	5	54	3	29	9	26	1	15	1	16	5	19	1	8	10	14	
34	56	4	23	6	21	1	59	7	17	5	38	14	49	6	14		4		37		41	8	59	2	61	4	33	1	6		12	8	22	
35	18		11	3	23	1	13	3	6	1	5	7	33	7					10		15	6	5	5	22	1	31	0			33	3	15	
36	3		58	1	18	4	22	3		1		1	4	1							2	7	19	3	11	4	20	7			15	1	6	
37			47	1	47		49	7														9	22	3	11	3	9	2			3		3	
38					33		37	8														11	37	6	43	2	30	2					7	
39					3		35	5														8	57	7	9	1	2	2					13	
40							2															2	2	6		3		4					9	
41																								2		1		1						
42																										4								
43																										5								
44																										2								
45																										2								
46																										2								

the results of such an analysis, the set of August/September length classes effectively considered is reduced to $\{L_j,\ j = m_1...m_2,^*\}$, where m_1 and m_2 are chosen such that they are the largest and smallest values respectively such that

$$\sum_{i \le m_1} y_i \le 5,$$

$$y_{m_2} \neq 0.$$

The trout in $(L_i,\ i \le m_1)$ are all partitioned into L_{m_1}. Thus, L_{m_1} contains the lower tail of the length distribution at t_2 comprising at least five trout. L_{m_2} contains the upper tail and is non-empty.

The construction of m_1 and m_2 reflects the different natures of the two tails at t_2. The small trout often have a wide range of lengths, so that the lower tail consists of length classes containing one or two trout with empty length classes between; these trout are grouped together to reduce the effect of their sparse allocation to length classes. The large trout have a narrow range of lengths and are partitioned into a few adjacent length classes; only the empty length classes that would contain very large fish need to be combined with the length class containing the largest observed trout. Table 9.2 shows the adjusted allocation of trout to the length classes.

9.3.6 Transition Probabilities

Assume that the transitions of individual trout between May/June and August/September length classes are independent. Let

$$p_{ij} = \text{prob(trout is in } L_j \text{ at } t_2 \mid \text{trout is in } K_i \text{ at } t_1),$$

$$j = m_1...m_2, \quad i = 1...60,$$

TABLE 9.2

The adjusted allocation of trout to May/June and Aug/Sep length classes

	1967		1968		1969		1970		1971		1972		1973		1974		1975		1976		1977		1978		1979		1980		1981		1982		1983	
i	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L	K	L
11													5																			
12								9					1													6						
13								2					2													2						
14								3					2							9				7		1						
15				5				2					1							6				1		1						
16		7		1		6		1					0							2				2		2						
17		6		3		0		2				5	2							2				1	6	0		6				5
18		4		1		0		4		6		3	8				5			4				0	1	0		1				1
19		4		1		0		2		3		6	5		5		0			1				5	4	0		4				0
20		3		1		6		2		4		3	8		0		7			3				4		1		1				5
21	2	7	2	7	3	2		3		9	4	6	4		3		0			7			5	2		0	3	8		6	5	1
22	1	12	10	6	9	6		5		3	9	11	6		1		0			2			2	9	3	1	14	2		5	10	4
23	7	5	8	1	11	2	1	4	4	9	26	5	23		2		5		7	4			9	1	27		8	6		2	9	7
24	2	8	4	2	7	2	4	3	1	10	10	5	6		0		5		8	11			12	2	23		4	6		6	15	3
25	7	7	12	9	4	4	31	2	6	3	15	3	4		9		9		19	3	1		10	2	18		15	11		5	29	2
26	21	5	29	1	12	5	23	3	12	4	38	8	11		1		5	13	12	8	3	7	2	2	9		19	5		2	49	1
27	6	5	6	8	10	7	14	4	17	9	51	3	63	15	2		8	5	24	8	3	3	4	6	23		23	2		6	34	1
28	75	4	36	6	33	8	12	4	56	7	39	9	24	6	9		8	2	45	7	3	9	11	4	8		4	2		2	51	1
29	36	3	35	6	23	4	39	5	48	10	63	6	52	2	23	8	8	7	47	3	4	4	12	5	14		14	2		4	43	1
30	18	5	11	3	20	3	48	6	17	8	9	6	36	3	8	12	5	6	18	2	7	3	8	4	5		5	2		0	39	0
31	45	0	24	2	15	3	12	5	55	8	41	4	55	11	11	12	7	7	56	4	5	4	10	8	19		19	4		0	19	0
32	27	16	58	8	45	5	23	9	43	3	23	4	29	4	17	1	9	5	17	4	11	3	8	2	21		16	3		5	33	5
33	22	11	35	7	66	5	37	2	26	7	68	3	13	4	19		14	5	54	9	36	11	26	1	15		16	1		0	14	4
34	56	4	23	6	21	1	59	7	17	5	38	14	49	6	14		4		37	2	29	8	59	2	61		33	1		7	22	2
35	18		11	3	23	1	13	3	6	1	5	7	33	7					10	5	41	6	5	5	22		31	0		2	15	2
36	3		58	1	18	4	22	3		1		1	4	1						2	15	7	19	3	11		20	7		2	6	1
37			47	1	47		49	7														9	22	6	11		9	2		2	3	0
38					33		37	8														11	37	3	43		30	2		4	7	5
39					3		35	5														8	57	7	9		2	4		1	13	4
40							2															2	2	6		1					9	
41																								2		4						
42																										1						
43																										4						
44																										5						
45																										2						
46																										2						

p_{i*} = prob(trout dies between t_1 and t_2 | trout is in K_i at t_1),

= prob(trout is in L_* at t_2 | trout is in K_i at t_1), $\quad i = 1...60.$

The distribution of the x_i trout in K_i at t_1 into the August/September length classes $\{L_j,\ j = m_1...m_2,*\}$ is a multinomial $\mathrm{Mn}(x_i;\ p_{im_1},...,p_{im_2},p_{i*})$. Thus, the distribution of trout at t_2 is given by the sum of multinomials

$$\sum_{i=1}^{60} \mathrm{Mn}(x_i;\ p_{im_1},\ldots,p_{im_2},p_{i*}).$$

Since none of the trout are marked, the available data effectively consist of the margins of a two way contingency table. The likelihood of the observed August/September length class distribution is

$$L(y \mid x, p) = \sum_{A} \prod_{i=1}^{60} \frac{x_i!}{z_{im_1}!\ldots z_{im_2}!\, z_{i*}!}\ p_{im_1}^{z_{im_1}} \cdots p_{im_2}^{z_{im_2}} p_{i*}^{z_{i*}},$$

where

$$A = \{z_{ij} : z_{ij} \text{ integer},\ \sum_{i=1}^{60} z_{ij} = y_j, j = m_1 \ldots m_2, *$$

$$\sum_{j=m_1}^{m_2} z_{ij} + z_{i*} = x_i,\ i = 1...60\}$$

The likelihood considers every way in which the population at t_1 can transfer to the population at t_2. The variables $\{z_{ij}\}$ denote the number of trout that move from

K_i to L_j and the constraints in the set A ensure that the correct number of trout belong to $\{K_i\}$ and $\{L_j\}$. However, the number of possible values of $\{z_{ij}\}$ is far too large for the likelihood to be of use.

It is much simpler to compare the expected and observed distributions of trout in the August/September length classes. The expected number of trout in L_j is

$$E_j = \sum_{i=1}^{\infty} x_i p_{ij}, \qquad j = m_1 \ldots m_2, *.$$

Two convenient measures by which sets of transition probabilities can be compared are a chi-squared measure

$$\chi^2(p) = \sum_{j=m_1}^{m_2} \frac{(E_j - y_j)^2}{E_j} + \frac{(E_\cdot - y_\cdot)^2}{E_\cdot} \tag{9.9}$$

and an information measure

$$I(p) = 2\{\sum_{j=m_1}^{m_2} y_j \log(y_j/E_j) + y_\cdot \log(y_\cdot/E_\cdot)\}. \tag{9.10}$$

The transition probabilities can be estimated by minimising either χ^2 (p) or $I(p)$ with respect to the probabilities.

First, however, constraints are imposed on the transition probabilities to reduce the number of parameters that have to be estimated to a useful and manageable quantity. An initial simplification is to write the probabilities as the product of a "survival" term and a "transition given survival" term

$$p_{ij} = s_i q_{ij}, \quad i = 1...60, \quad j = m_1...m_2,$$

$$p_{i\cdot} = (1 - s_i), \quad i = 1...60,$$

where

$$\sum_{j=m_1}^{m_2} q_{ij} = 1, \quad i = 1...60.$$

Thus, s_i is the probability that a trout initially in K_i survives from t_1 to t_2 and q_{ij} is the probability that a trout transfers from K_i to L_j given that it survives from t_1 to t_2.

Biological arguments now suggest parametric forms for the probabilities (s_i) and (q_{ij}). Three main criteria are always sought.

i) Trout in neighbouring length classes at t_1 have similar survival probabilities.

ii) The probabilities q_{ij}, $j = m_1...m_2$, are "unimodal" for each i. A trout from K_i that survives from t_1 to t_2 is most likely to go to a modal length class $L_{\hat{j}(i)}$ say and is progressively less likely to go to L_j the "further" L_j is from $L_{\hat{j}(i)}$ (Fig. 9.5). Formally, as j increases from m_1+1 to m_2-1, q_{ij} increases monotonically to some maximum value and decreases monotonically thereafter. (This definition of unimodality ignores q_{im_1} and q_{im_2} since these

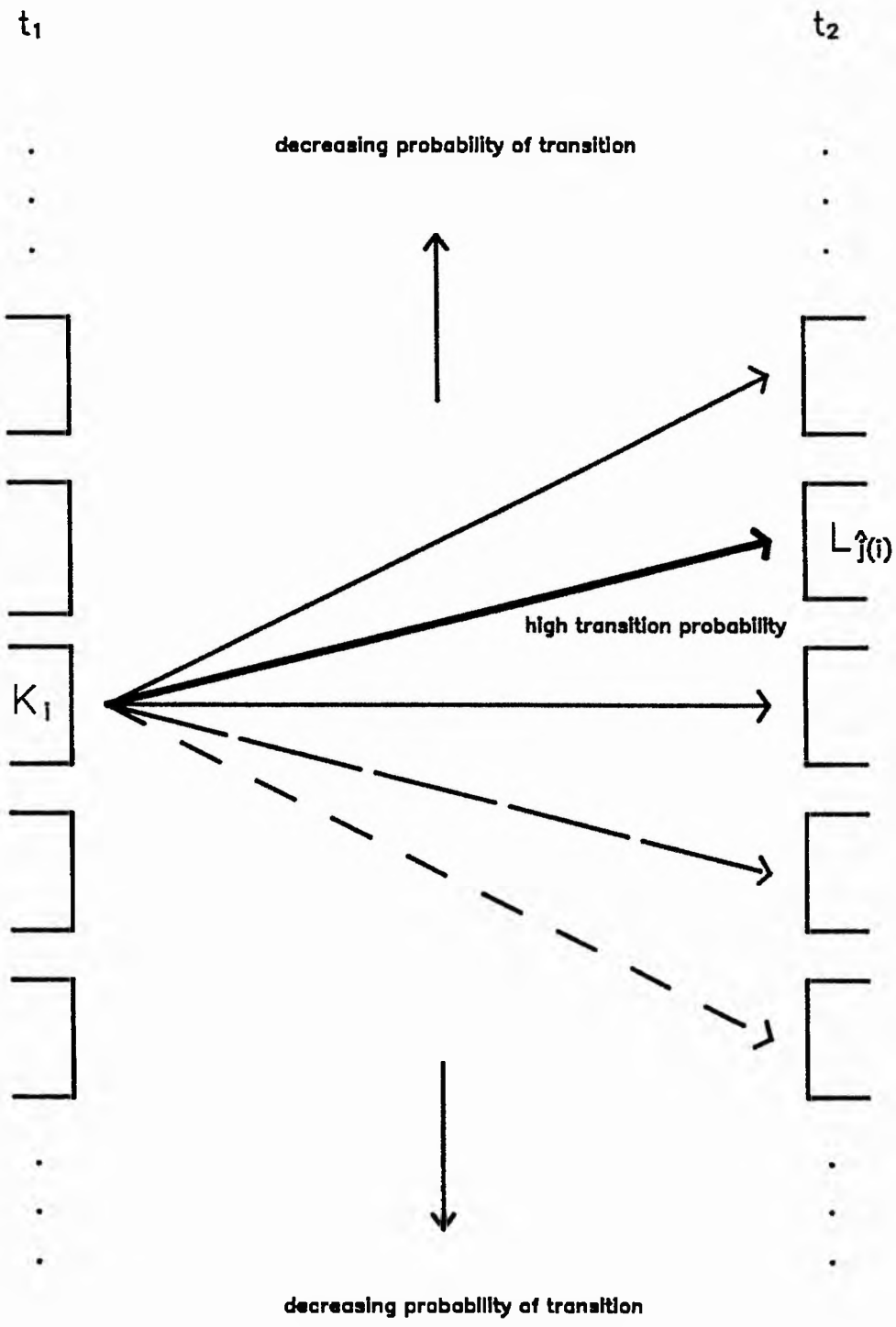


Fig. 9.5 Unimodal transition probabilities.

are the probabilities that a trout in K_i at t_1 has a length in the lower or upper tail of the length distribution at t_2, so that it is possible that $q_{im_1} \geq q_{i,m_1+1}$ and $q_{im_2} \geq q_{i,m_2-1}$).

iii) If two trout survive from t_1 to t_2, then on average, the larger trout at t_1 is still the larger at t_2.

The next section describes various parameterisations of $\{s_i\}$ and $\{q_{ij}\}$. Parameter estimates for the years 1967 to 1983 are given for each model and the ability of each model to reflect the observed data is discussed.

9.4 MODEL PARAMETERISATIONS AND RESULTS

9.4.1 Definitions

9.4.1.1 10th and 90th percentile lengths

The 10th and 90th percentile lengths are useful measures that describe the lengths of a relatively small trout at t_1 and a relatively large trout at t_1 respectively. The 10th percentile length l_L is the smallest length (to the nearest mm) such that

$$\sum_{i \leq l_L} x_i \geq \frac{1}{10} R_1.$$

Similarly, the 90th percentile length l_U is the smallest length (to the nearest mm) such that

$$\sum_{i \leq l_U} x_i \geq \frac{9}{10} R_1.$$

For example, in 1967, l_L = 26 mm and l_U = 34 mm. Any structure or parameter with a suffix L or U relates to trout of lengths l_L or l_U; for example, s_L is the survival probability of a trout of length l_L (in K_L) at t_1.

9.4.1.2 Relative transition mean

Consider a trout initially in K_i that survives from t_1 to t_2. The mean length of the trout at t_2 is approximately

$$\sum_{j=m_1}^{m_2} f(j)q_{ij}.$$

(This is an approximation because

i) $f(j)$ is not the mid-point of $f(j - 0.5)$ and $f(j + 0.5)$,

ii) there are slight errors introduced by the construction of the tail end classes L_{m_1} and L_{m_2}).

The relative transition mean of the trout is

$$\mu_i = \sum_{j=m_1}^{m_2} jq_{ij} - i.$$

This gives the mean position of the transition class to which the trout transfers relative to L_i (, which is the transition class to which a trout would transfer if there were no variation and if the trout were feeding on maximum rations). For example, if $\mu_i = 0$, then on average, the trout moves to L_i and is growing as predicted by the growth function f. If $\mu_i = -1$, then on average, the trout moves to L_{i-1}; the trout does not grow as well as predicted, which is consistent with feeding on less than

maximum rations. Values of $\mu_i > 0$ indicate that the trout are growing larger than predicted and that there are slight inadequacies in the growth model (9.7). Only the relative transition mean is considered and it is abbreviated to the transition mean.

9.4.1.3 Transition variance

The transition variance of a trout from K_i that survives from t_1 to t_2 is

$$\sigma_i^2 = \sum_{j=m_1}^{m_2} j^2 q_{ij} - \mu_i^2.$$

A minimum plausible value of σ_i^2 is obtained by a simulation. Equations (9.6), (9.7) and (9.8) have stochastic components which are ignored in the construction of $f(.)$. By generating these stochastic components, the growth of 1,000 trout of length i mm at t_1 is followed to obtain a distribution of the corresponding length at t_2. This distribution suggests that trout transfer to length classes up to two away from the transition mean with a relatively high probability and that the transition standard deviation is approximately one transition class. Since the simulation does not include errors in measuring lengths at t_1 and t_2, nor the errors in the random partitioning of trout into length classes at t_2, nor the variation in ration size between trout, the transition variance is likely to be greater than 1. Hence, estimated values of σ_i^2 below 1 must be considered unrealistic.

9.4.1.4 Approximate means and variances

The transition probabilities $\{q_{ij}\}$ are generally written as the discrete approximation to a continuous distribution. For example, a discrete approximation to the Normal distribution is

$$q_{ij} = \frac{k_i}{\tau_i} \exp\left\{-\frac{1}{2\tau_i^2}(j-i-\eta_i)^2\right\}, \qquad j = m_1+1,\ldots,m_2-1,$$

$$q_{im_1} = \frac{k_i}{\tau_i} \sum_{j=-\infty}^{m_1} \exp\left\{-\frac{1}{2\tau_i^2}(j-i-\eta_i)^2\right\},$$

$$q_{im_2} = \frac{k_i}{\tau_i} \sum_{j=m_2}^{\infty} \exp\left\{-\frac{1}{2\tau_i^2}(j-i-\eta_i)^2\right\},$$

$$i = 1\ldots60. \qquad (9.11)$$

The probabilities q_{im_1}, q_{im_2} are the sums of the lower and upper tails of the discretised distribution respectively and the normalising constant k_i is given by

$$k_i = \left(\frac{1}{\tau_i} \sum_{j=-\infty}^{\infty} \exp\left\{-\frac{1}{2\tau_i^2}(j-i-\eta_i)^2\right\}\right)^{-1}, \qquad i = 1\ldots60.$$

For biologically realistic values of τ_i^2 (ie $\tau_i^2 \geq 1.0$), the discrete normal distribution (9.11) is very similar to the grouped normal distribution; see, for example, the qualitative comparison of the two probability distributions for different values of $\eta_i, \tau_i^2 = 1.0$ (Table 9.3). However, the discrete distribution has computational advantages over the grouped distribution. The transition and survival probabilities are estimated by minimising the chi-squared measure (9.9) or the information measure (9.10). The goodness of fit surfaces often prove to be relatively flat. Minimisation routines are therefore more efficient if derivatives can be provided and these are much easier to compute for the discrete distribution (9.11) than for the grouped distribution.

The transition mean and variance μ_i and σ_i^2 can be approximately related to the parameters of the discrete distribution (in this case η_i, τ_i^2) provided these parameters

TABLE 9.3

Properties of the discrete normal distribution

	Comparison of the discrete and grouped normal distribution			
	$\eta_i = 0,\ \tau_i^2 = 1.0$		$\eta_i = -0.2,\ \tau_i^2 = 1.0$	
	Transition probabilities		Transition probabilities	
	Discrete	Grouped	Discrete	Grouped
$q_{i,i-4}$	0.00013	0.00023	0.00029	0.00047
$q_{i,i-3}$	0.00443	0.00598	0.00792	0.01024
$q_{i,i-2}$	0.05399	0.06060	0.07895	0.08608
$q_{i,i-1}$	0.24197	0.24173	0.28969	0.28529
$q_{i,i}$	0.39894	0.38293	0.39104	0.37595
$q_{i,i+1}$	0.24197	0.24173	0.19419	0.19740
$q_{i,i+2}$	0.05399	0.06060	0.03547	0.04110
$q_{i,i+3}$	0.00443	0.00598	0.00238	0.00336
$q_{i,i+4}$	0.00013	0.00023	0.00006	0.00011

Comparison of (η_i, τ_i^2) and (μ_i, σ_i^2)		
$\tau_i^2 = 1.0$		
η_i	μ_i	σ_i^2
0.00	0.00000000	0.99999976
-0.20	-0.20000002	0.99999994
-0.50	-0.50000000	1.00000012

take sensible values. For example, for realistic values of τ_i^2, the integral to obtain the mean of a continuous normal distribution can be approximated by a summation, so that

$$i + \eta_i \cong \sum_{j=-\infty}^{\infty} \frac{jk_i}{\tau_i} \exp\left\{-\frac{1}{2\tau_i^2}(j-i-\eta_i)^2\right\}$$

If the tail transition probabilities q_{im_1} and q_{im_2} are small, then

$$\eta_i \cong jq_{ij} - i = \mu_i.$$

Similarly, $\tau_i^2 \cong \sigma_i^2$. The order of the remainder term is very small (Table 9.3). Thus, constraints on the parameters can be immediately interpreted in terms of approximate constraints on the transition mean and variance; for example, if η_i is constant over i, the transition mean is approximately constant. This approximate relationship will be implicitly assumed in future.

9.4.1.5 Goodness of fit and degrees of freedom

The degrees of freedom for a given model is the difference between the number of August/September length classes (excluding L_*) and the number of parameters in the model. A rough guide to the goodness of fit of a model to all 17 years data is obtained by comparing the sum of the minimised chi-squared measures ($\sum \hat{\chi}^2$) or the sum of the minimised information measures ($\sum \hat{I}$) to a chi-squared distribution with degrees of freedom given by the sum of the individual degrees of freedom.

9.4.2 $\{s_i\}$ Logit

Except for model M8, all the models assume that the survival probabilities (s_i) are of the form

$$s_i = \exp(\theta + \phi i)/\{1 + \exp(\theta + \phi i)\}, \qquad i = 1...60,$$

where θ, ϕ are two parameters to be estimated. The probability of survival either increases or decreases monotonically with i or is constant. The computation involved in estimating the parameters is simplified by transforming the parameters to

$$g_L = \theta + \phi l_L,$$

$$g_U = \theta + \phi l_U.$$

Thus

$$s_L = \exp(g_L)/\{1 + \exp(g_L)\},$$

$$s_U = \exp(g_U)/\{1 + \exp(g_U)\}.$$

9.4.3 $\{q_{ij}\}$ Discrete Normal: Constant Mean and Variance

Model M1:

Let the transition probabilities $\{q_{ij}\}$ be given by the discrete normal distribution (9.11) and assume that $\eta_i = \eta$ and $\tau_i = \tau$ are constant for all i. Thus, two parameters η and τ are to be estimated. The transition mean and variance are the same for trout from all the length classes $\{K_i\}$, and the distribution of transitions from K_i is approximately symmetric about η (Fig. 9.6).

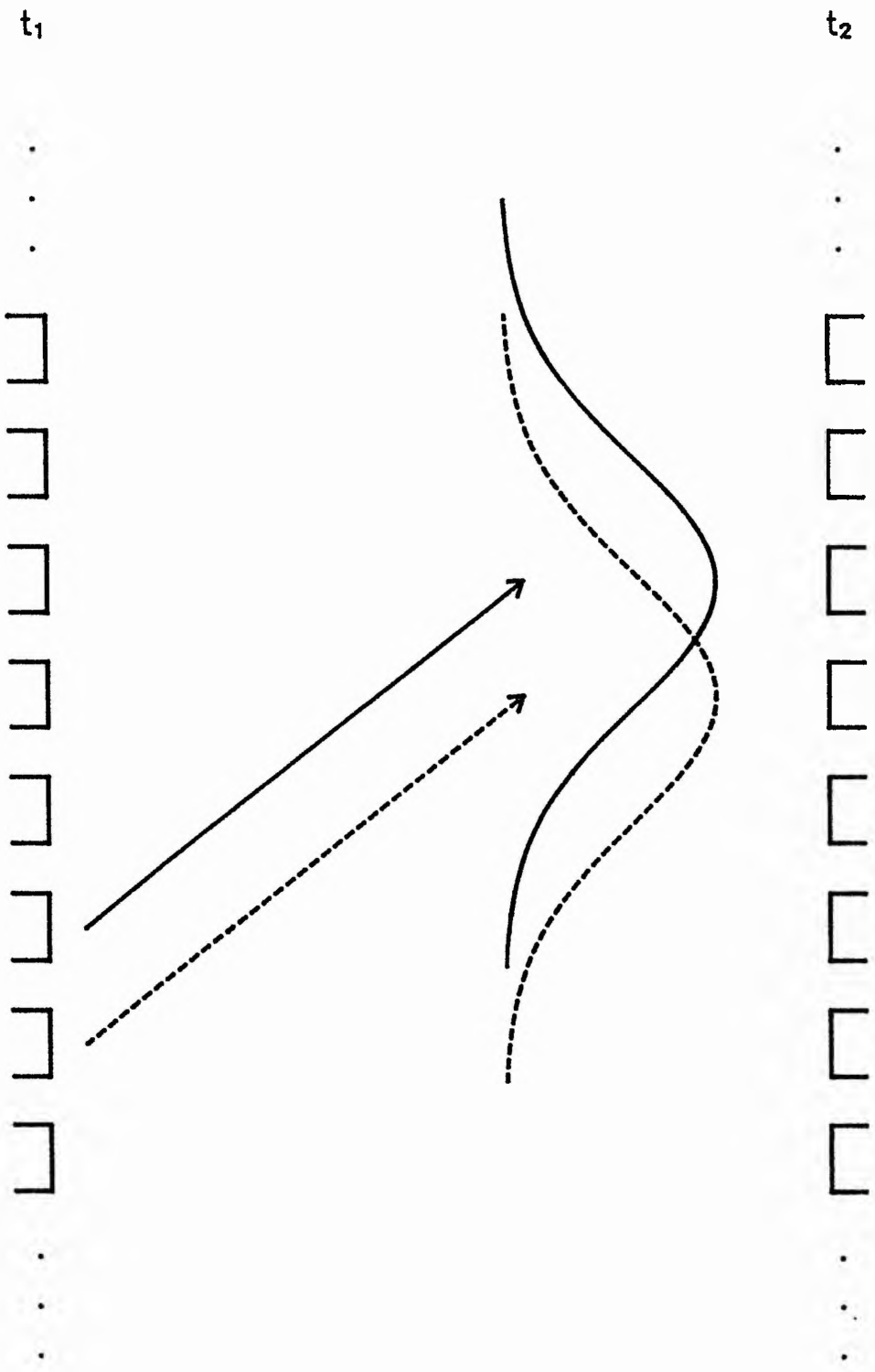


Fig. 9.6 The transition probabilities $\{q_{ij}\}$ given by the discrete normal distribution with constant mean and constant variance.

The parameter estimates obtained by minimising the chi-squared measure are given in Table 9.4. In general, the estimates give plausible values for the parameters. Except in 1977 and 1979, the transition mean is negative, which is consistent with growth on less than maximum rations. The transition standard deviation is not too small, except in 1983 which is a year badly affected by drought; however in some years τ is so large that trout transfer to a wide range of length classes with a very similar probability. The survival probability of a large trout s_U is nearly always greater than the survival probability of a small trout s_L.

The fit of the model is not good ($\Sigma\hat{\chi}^2 = 589.27$ on 282 d.f.). This is because the model can not adequately explain the different forms of the tails of the length distribution at t_2. There are two areas of conflicting requirements:

i) a small transition variance is needed to account for the tight grouping of the upper tail and a large transition variance to account for the spread of the lower tail.

ii) a high transition mean is needed to account for the good growth of large trout at t_1 and a low transition mean to account for the poor growth of small trout at t_2.

The parameter estimates obtained by minimising the information measure are given in Table 9.5. The two measures give very similar results. In fact, both measures give similar parameter estimates for each of the models that follow; for simplicity, only the results obtained using the chi-squared measure are reported here.

TABLE 9.4

Parameter estimates for model M1: chi-squared measure

Year	E	R_1	R_2	$\hat{\chi}^2$	df	η	τ	g_L	g_U	l_L	l_U	s_L	s_U
1980	722	286	61	40.29	25	-4.53	8.52	-1.53	-0.62	24	38	0.18	0.35
1979	1034	334	73	36.18	26	1.08	7.68	-0.83	-1.27	24	38	0.30	0.22
1969	1756	403	74	20.63	17	-6.45	4.28	-1.56	-1.21	26	37	0.17	0.23
1978	2068	429	105	48.59	24	-4.53	7.49	-1.28	-0.56	27	39	0.22	0.36
1968	2272	409	92	31.49	18	-5.01	4.21	-1.10	-1.06	26	37	0.25	0.26
1983	2478	425	43	16.61	12	-8.09	0.86	-2.85	-1.32	25	35	0.05	0.21
1970	2584	460	115	36.14	24	-6.22	8.54	-1.84	-0.35	26	38	0.14	0.41
1972	2888	439	112	42.07	15	-1.69	4.89	-0.84	-0.96	24	33	0.30	0.28
1973	3610	412	119	42.01	21	-5.43	6.11	-1.32	-0.29	25	34	0.21	0.43
1976	4234	354	91	49.34	16	-6.32	6.26	-1.25	-0.44	26	34	0.22	0.39
1971	4644	308	110	21.13	14	-3.96	4.40	-0.91	-0.11	27	33	0.29	0.47
1967	4652	346	116	62.34	15	-4.70	5.39	-1.03	0.01	26	34	0.26	0.50
1977	6300	241	99	22.99	11	1.77	3.85	-0.59	0.05	29	34	0.36	0.51
1975	7334	182	88	33.87	12	-3.73	4.12	-0.63	1.08	27	33	0.35	0.75
1981	7646	178	87	36.50	12	-4.09	4.20	0.09	0.18	26	33	0.52	0.54
1974	7958	132	74	21.74	9	-4.36	3.86	-0.41	1.64	29	34	0.40	0.84
1982	7958	158	81	27.35	11	-4.23	4.12	-0.17	0.66	30	36	0.46	0.66
				589.27	282								

TABLE 9.5

Parameter estimates for model M1: information measure

Year	E	R_1	R_2	$\hat{I}$	df	η	τ	g_L	g_U	l_L	l_U	s_L	s_U
1980	722	286	61	48.51	25	-4.97	9.29	-1.85	-0.88	24	38	0.14	0.29
1979	1034	334	73	37.20	26	-0.48	7.76	-1.66	-0.95	24	38	0.16	0.28
1969	1756	403	74	25.82	17	-6.53	4.30	-1.91	-1.20	26	37	0.13	0.23
1978	2068	429	105	51.65	24	-5.57	7.24	-1.62	-0.65	27	39	0.17	0.34
1968	2272	409	92	36.92	18	-5.25	4.22	-1.32	-1.15	26	37	0.21	0.24
1983	2478	425	43	21.67	12	-7.92	0.84	-2.99	-1.53	25	35	0.05	0.18
1970	2584	460	115	34.08	24	-7.02	8.16	-1.93	-0.54	26	38	0.13	0.37
1972	2888	439	112	34.04	15	-1.79	4.81	-0.80	-1.30	24	33	0.31	0.21
1973	3610	412	119	43.50	21	-5.48	6.31	-1.51	-0.43	25	34	0.18	0.39
1976	4234	354	91	42.46	16	-7.29	5.99	-1.54	-0.64	26	34	0.18	0.35
1971	4644	308	110	22.87	14	-4.22	4.33	-1.11	-0.11	27	33	0.25	0.47
1967	4652	346	116	58.21	15	-5.32	5.14	-0.93	-0.48	26	34	0.28	0.38
1977	6300	241	99	24.16	11	1.50	3.64	-0.72	-0.05	29	34	0.33	0.49
1975	7334	182	88	42.50	12	-3.74	3.98	-1.08	1.23	27	33	0.25	0.77
1981	7646	178	87	33.00	12	-4.65	3.93	-0.16	0.05	26	33	0.46	0.51
1974	7958	132	74	27.22	9	-4.14	3.74	-0.63	1.61	29	34	0.35	0.83
1982	7958	158	81	26.84	11	-4.47	4.08	-0.35	0.54	30	36	0.41	0.63
				610.65	282								

9.4.4 $\{q_{ij}\}$ Discrete Extreme Value: Constant Mean and Variance

Model M2:

The normal model gives a symmetric distribution of transitions from K_i. The form of the tails of the length distribution at t_2 suggests that a negatively skewed model might be more appropriate. This can be interpreted as follows. Most of the trout that survive from K_i grow to roughly the same size at t_2; these are the "fittest for survival" and few trout from K_i grow to a larger size. However, quite a number of trout that are "less fit" still survive; these grow smaller than the main group, giving a long lower tail. A convenient negatively skewed distribution is an extreme value distribution; a discrete approximation is:

$$q_{ij} = \frac{k_i}{\gamma_i} \exp\left\{\frac{j-i-\delta_i}{\gamma_i} - \exp\left\{\frac{j-i-\delta_i}{\gamma_i}\right\}\right\}, \qquad j = m_1+1,\ldots,m_2-1,$$

$$q_{im_1} = \frac{k_i}{\gamma_i} \sum_{j=-\infty}^{m_1} \exp\left\{\frac{j-i-\delta_i}{\gamma_i} - \exp\left\{\frac{j-i-\delta_i}{\gamma_i}\right\}\right\},$$

$$q_{im_2} = \frac{k_i}{\gamma_i} \sum_{j=m_2}^{\infty} \exp\left\{\frac{j-i-\delta_i}{\gamma_i} - \exp\left\{\frac{j-i-\delta_i}{\gamma_i}\right\}\right\},$$

$$i = 1\ldots60.$$

Assuming that $\gamma_i = \gamma$ and $\delta_i = \delta$ are constant over i, the transition mean and variance are constant and given by

$$\mu = \delta - \gamma\upsilon,$$

$$\sigma^2 = (\pi\gamma)^2/6,$$

where ν is Euler's constant (≈ 0.58). The skewness of the distribution is approximately -1.3 (Fig. 9.7).

Parameter estimates are given in Table 9.6. The overall fit of the model is still not good ($\Sigma\hat{X}^2 = 525.07$ on 282 d.f.), although it is better than that of model M1. There are numerical difficulties in estimating the parameters for six of the years; these occur when the estimates are in regions of the parameter space that are biologically implausible. For example, in 1983, the estimate of the transition variance is very small. Also, in 1970, 1974, 1976, 1980, and 1982, the minimisation routine is halted as the estimates of the survival parameters g_L and g_U are tending to either $+\infty$ or $-\infty$. The estimates of g_L and g_U are negatively correlated since the model has to account for an observed number of trout that survive; if g_L decreases, the survival probability of small trout decreases so that the survival probability of large trout and hence g_U must increase to compensate. Thus, in these five years, the estimated distribution of the survival probabilities is a step function; either all the trout below a certain length die and all trout above it survive or vice versa.

The extreme value model gives a first indication that the chi-squared goodness of fit surface can be very flat and that it is important to try a number of different initial parameter estimates in any minimisation routine. For example, in 1974, the parameter estimates

$$\gamma = -1.06, \quad \delta = 2.96, \quad g_L = 9.27, \quad g_U = -10.0,$$

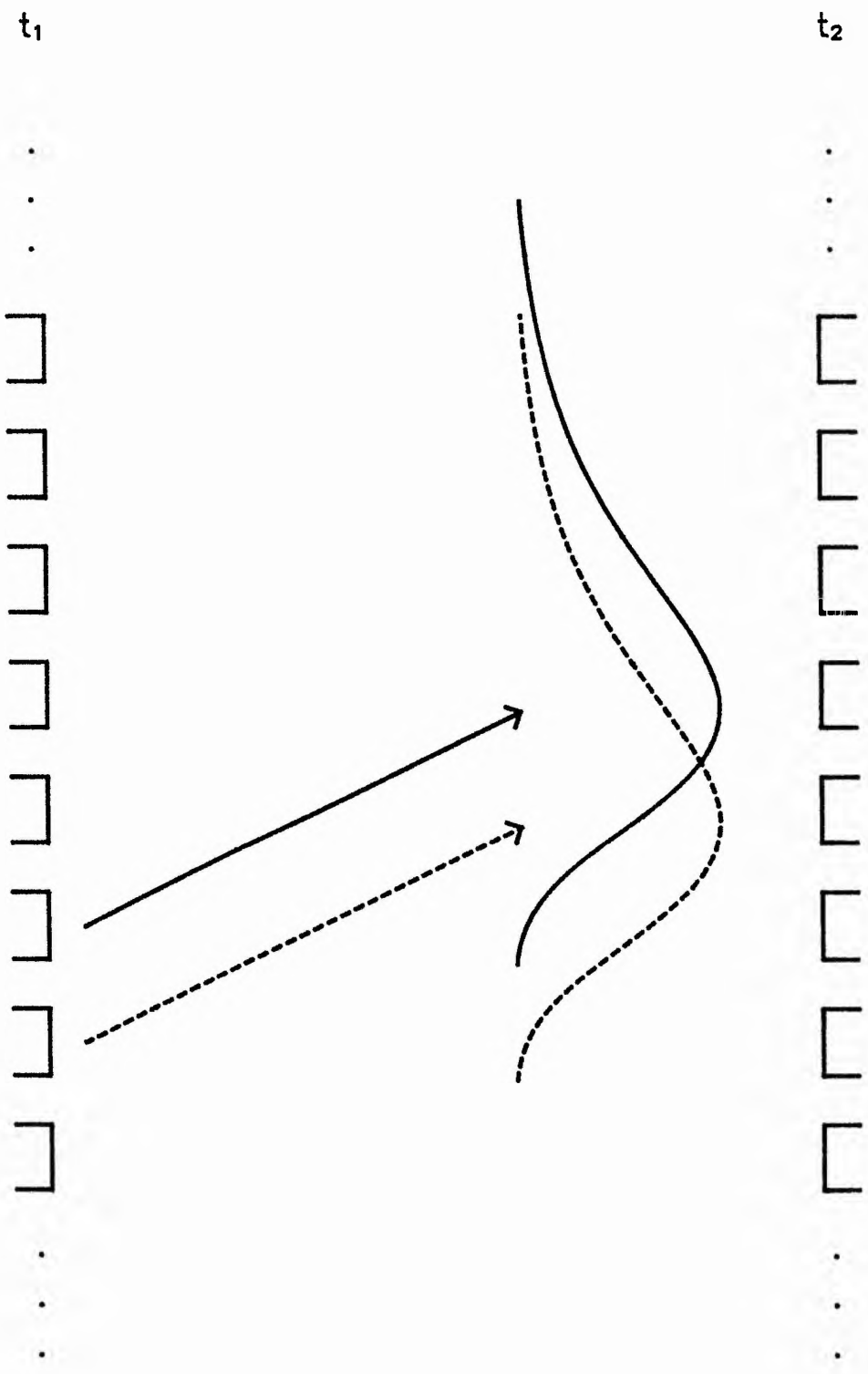


Fig. 9.7 The transition probabilities $\{q_{ij}\}$ given by the discrete extreme value distribution with constant mean and constant variance.

TABLE 9.6

Parameter estimates for model M2

Year	E	R_1	R_2	$\hat{X}^2$	df	γ	δ	μ	σ	g_L	g_U	l_L	l_U	s_L	s_U	
1980	722	286	61	34.33	25	-4.74	8.10	-9.44	10.38	-10.00	2.34	24	38	0.00	0.91	**
1979	1034	334	73	34.53	26	3.59	6.22	-0.02	7.98	-0.99	-1.13	24	38	0.27	0.24	
1969	1756	403	74	20.07	17	-2.07	3.00	-3.82	3.85	-0.59	-2.20	26	37	0.36	0.10	
1978	2068	429	105	45.33	24	-2.34	6.22	-5.95	7.98	-1.60	-0.30	27	39	0.17	0.43	
1968	2272	409	92	27.95	18	-3.61	3.36	-5.56	4.30	-1.26	-0.94	26	37	0.22	0.28	
1983	2478	425	43	16.96	12	-7.74	0.80	-8.20	1.03	-2.89	-1.29	25	35	0.05	0.22	**
1970	2584	460	115	24.25	24	-6.08	7.70	-10.54	9.87	-10.00	1.85	26	38	0.00	0.86	**
1972	2888	439	112	42.44	15	-0.65	0.41	-3.01	5.20	-1.35	-0.59	24	33	0.21	0.36	
1973	3610	412	119	38.51	21	-2.62	4.97	-5.51	6.38	-1.14	-0.43	25	34	0.24	0.39	
1976	4234	354	91	44.61	16	-6.40	5.90	-9.82	7.57	-10.00	3.13	26	34	0.00	0.96	**
1971	4644	308	110	22.17	14	-2.04	3.76	-4.22	4.83	-0.79	-0.21	27	33	0.31	0.45	
1967	4652	346	116	58.39	15	-3.81	4.67	-6.52	5.99	-2.41	0.90	26	34	0.08	0.71	
1977	6300	241	99	18.22	11	2.36	3.31	0.44	4.24	-2.30	1.26	29	34	0.09	0.78	
1975	7334	182	88	31.72	12	-2.27	3.40	-4.25	4.36	-0.97	1.52	27	33	0.28	0.82	
1981	7646	178	87	31.02	12	-2.70	3.29	-4.61	4.22	-0.23	0.38	26	33	0.44	0.59	
1974	7958	132	74	13.58	9	-3.47	3.02	-5.22	3.87	-5.46	10.00	29	34	0.00	1.00	**
1982	7958	158	81	20.99	11	-0.60	3.57	-2.66	4.57	9.31	-10.00	30	36	1.00	0.00	**
				525.07	282											

** indicates serious difficulty in obtaining parameter estimates

give a chi-squared value of 13.67, whilst the parameter values

$$\gamma = 3.47, \quad \delta = 3.02, \quad g_L = -5.25, \quad g_U = 10.0,$$

virtually reverse the survival probabilities and give a similar chi-squared value of 13.58.

The extreme value model is not developed further because of these problems in obtaining and interpreting the parameter estimates.

9.4.5 $\{q_{ij}\}$ Discrete Normal: More Complicated Constraints

The discrete normal model (9.11) is now considered in more detail. Three groups of constraints are imposed on the parameters $\{\eta_i\}$ and $\{\tau_i^2\}$; the biological interpretation and the fit of each model are discussed in turn.

9.4.5.1 $\{\eta_i\}$ varies linearly

Model M3:

$$\eta_i = \theta + \phi i, \qquad i = 41 \ldots 60.$$
$$\tau_i = \tau,$$

The transition mean varies linearly with the trout length at t_1 and the transition variance is constant. The parameters are transformed to

$$\eta_L = \theta + \phi l_L,$$

$$\eta_U = \theta + \phi l_U,$$

which are the transition means of a small and a large trout at t_1 respectively. Parameter estimates are given in Table 9.7.

The overall fit of this model is much better than that of model M1 ($\sum \hat{X}^2 = 369.42$ on 265 d.f.). The transition mean of small trout η_L is always less than the transition mean of large trout η_U; an interpretation is that the large trout can compete for food better than the small trout and hence grow larger relative to their original size. The very large negative values of η_L (for example, $\eta_L = -19.79$ in 1980) do not indicate a limitation of the model since they always occur when the survival probability of small trout s_L is low; few large negative transitions are predicted since most of the small trout at t_1 die between t_1 and t_2. Unfortunately, there is no obvious pattern in the survival probabilities, the transition variances are often small and the chi-squared goodness of fit surface is very flat.

9.4.5.2 {τ_i} inversely proportional to length

Model M4:

$$\tau_i = k/i,$$

$$i = 1...60.$$

$$\eta_i = \eta,$$

The transition standard deviation is inversely proportional to the length of the trout at t_1 and the transition mean is constant. It is assumed that, as length at t_1 decreases, the ability to compete for food becomes more variable so that transition variance

TABLE 9.7

Parameter estimates for model M3

Year	E	R_1	R_2	$\hat{X}^2$	df	η_L	η_U	τ	g_L	g_U	l_L	l_U	s_L	s_U	
1980	722	286	61	32.46	24	-19.79	-0.55	2.97	-2.47	-0.12	24	38	0.08	0.47	
1979	1034	334	73	31.37	25	-1.97	5.55	4.24	-0.68	-1.49	24	38	0.34	0.18	
1969	1756	403	74	20.35	16	-14.19	-7.85	4.01	-3.45	-0.34	26	37	0.03	0.42	
1978	2068	429	105	17.37	23	-8.13	0.52	0.83	-0.93	-1.19	27	39	0.28	0.23	
1968	2272	409	92	20.08	17	-8.55	-3.00	1.23	-1.28	-0.99	26	37	0.22	0.27	
1983	2478	425	43	11.80	11	-6.07	-2.43	0.31	-1.98	-2.16	25	35	0.12	0.10	
1970	2584	460	115	19.20	23	-21.61	-3.52	2.66	-2.61	-0.07	26	38	0.07	0.48	
1972	2888	439	112	29.54	14	-4.47	0.46	1.20	-0.72	-1.14	24	33	0.32	0.24	
1973	3610	412	119	33.16	20	-9.62	-1.88	1.26	-1.08	-0.52	25	34	0.25	0.37	
1976	4234	354	91	35.32	15	-8.60	-0.88	1.05	-0.48	-1.34	26	34	0.38	0.21	
1971	4644	308	110	9.89	13	-7.59	-1.65	0.64	-0.95	-0.16	27	33	0.28	0.46	
1967	4652	346	116	27.00	14	-8.05	-1.72	0.80	-0.50	-0.64	26	34	0.38	0.35	
1977	6300	241	99	8.36	10	-0.18	4.01	0.74	-0.24	-0.39	29	34	0.43	0.40	
1975	7334	182	88	24.86	11	-12.67	-1.31	0.87	-4.02	6.09	27	33	0.02	0.99	
1981	7646	178	87	16.29	11	-5.40	1.48	0.30	7.31	-3.63	26	33	0.99	0.03	**
1974	7958	132	74	12.81	8	-7.67	-2.40	0.30	-0.64	1.74	29	34	0.34	0.85	**
1982	7958	158	81	19.56	10	-12.37	-2.82	2.06	-3.74	5.07	30	36	0.02	0.99	
				369.42	265										

** indicates serious difficultly in obtaining parameter estimates

increases. It is important to note that the lengths of the trout at t_1 only vary between 21 mm and 40 mm. Hence, the transition variances of small and large trout are only allowed to vary by at most a factor of four. Parameter estimates are given in Table 9.8.

The overall fit of the model is similar to that of model M1 ($\Sigma\hat{\chi}^2$ = 561.00 on 282 d.f.). However, the estimates of the survival probabilities are very different; the survival probability of a large trout is greater than that of a small trout in only four years out of seventeen. This occurs because the "fittest" small trout can grow bigger relative to their original size than the "fittest" large trout. Also, in general, the transition variance is quite large. Therefore, if the survival probability of a small trout is high, the small trout at t_1 can "account" for all the small trout and many of the large trout at t_2. The rest of the large trout at t_2 come from the comparatively few large trout at t_1 that survive.

Model M5:

Greater flexibility is obtained by allowing the transition mean to vary linearly:

$$\tau_i = k/i,$$

$$\eta_i = \theta + \phi i,$$

$$i = 1...60.$$

Parameter estimates are given in Table 9.9.

The overall fit of the model ($\Sigma\chi^2$ = 366.15 on 265 d.f.) and the parameter estimates are very similar to those of model M3. There is no obvious pattern in the survival probabilities.

TABLE 9.8

Parameter estimates for model M4

Year	E	R_1	R_2	$\hat{X}^2$	df	η	k	τ_L	τ_U	g_L	g_U	l_L	l_U	s_L	s_U
1980	722	286	61	36.83	25	-2.60	234.7	9.78	6.18	-0.63	-1.46	24	38	0.35	0.19
1979	1034	334	73	35.30	26	3.72	210.4	8.77	5.54	-0.09	-2.50	24	38	0.52	0.08
1969	1756	403	74	20.43	17	-7.78	165.5	6.37	4.47	-2.28	-0.79	26	37	0.09	0.31
1978	2068	429	105	46.88	24	-2.26	219.8	8.14	5.64	-0.38	-1.65	27	39	0.40	0.16
1968	2272	409	92	29.31	18	-3.99	116.3	4.47	3.14	-0.69	-1.55	26	37	0.33	0.18
1983	2478	425	43	15.88	12	-7.99	31.7	1.27	0.91	-2.82	-1.35	25	35	0.06	0.21
1970	2584	460	115	31.17	24	-4.33	250.0	9.62	6.58	-0.75	-1.17	26	38	0.32	0.24
1972	2888	439	112	41.88	15	-0.78	134.1	5.59	4.06	-0.25	-1.48	24	33	0.44	0.19
1973	3610	412	119	39.41	21	-4.10	167.3	6.69	4.92	-0.48	-1.01	25	34	0.38	0.27
1976	4234	354	91	47.77	16	-4.99	175.4	6.75	5.16	-0.34	-1.38	26	34	0.42	0.20
1971	4644	308	110	20.80	14	-3.07	127.3	4.71	3.86	-0.13	-0.89	27	33	0.47	0.29
1967	4652	346	116	60.65	15	-3.66	152.9	5.88	4.50	-0.14	-0.75	26	34	0.46	0.32
1977	6300	241	99	21.49	11	2.20	116.7	4.02	3.43	-0.10	-0.41	29	34	0.47	0.40
1975	7334	182	88	32.75	12	-3.21	117.6	4.36	3.56	0.01	0.21	27	33	0.50	0.55
1981	7646	178	87	34.71	12	-3.63	118.3	4.55	3.58	0.81	-0.39	26	33	0.69	0.40
1974	7958	132	74	19.96	9	-4.06	114.5	3.95	3.37	0.08	0.79	29	34	0.52	0.69
1982	7958	158	81	25.78	11	-3.97	131.0	4.37	3.64	0.24	0.20	30	36	0.56	0.55
				561.00	282										

TABLE 9.9

Parameter estimates for model M5

Year	E	R_1	R_2	$\hat{X}^2$	df	η_L	η_U	k	τ_L	τ_U	g_L	g_U	l_L	l_U	s_L	s_U	
1980	722	286	61	32.65	24	-20.12	-0.59	113.4	4.72	2.98	-2.48	-0.12	24	38	0.08	0.47	
1979	1034	334	73	29.69	25	-1.53	5.86	113.4	4.72	2.98	-0.59	-1.62	24	38	0.36	0.17	
1969	1756	403	74	20.43	16	-8.59	-7.88	165.4	6.36	4.47	-2.49	-0.69	26	37	0.08	0.33	
1978	2068	429	105	16.84	23	-8.31	0.59	29.2	1.08	0.75	-0.95	-1.17	27	39	0.28	0.24	
1968	2272	409	92	20.89	17	-8.36	-2.94	43.2	1.66	1.17	-1.24	-1.03	26	37	0.23	0.26	
1983	2478	425	43	11.51	11	-6.04	-2.37	8.4	0.34	0.24	-1.98	-2.17	25	35	0.12	0.10	
1970	2584	460	115	18.80	23	-21.43	-3.43	94.0	3.62	2.47	-2.55	-0.09	26	38	0.07	0.47	
1972	2888	439	112	28.26	14	-4.39	0.53	35.9	1.50	1.09	-0.68	-1.18	24	33	0.34	0.24	
1973	3610	412	119	33.67	20	-9.71	-1.89	37.0	1.48	1.09	-1.10	-0.50	25	34	0.26	0.38	
1976	4234	354	91	35.28	15	-8.39	-0.99	35.3	1.36	1.04	-0.45	-1.38	26	34	0.39	0.20	
1971	4644	308	110	9.58	13	-7.60	-1.63	19.4	0.72	0.59	-0.95	-0.17	27	33	0.28	0.46	
1967	4652	346	116	24.89	14	-8.03	-1.69	23.0	0.88	0.68	-0.49	-0.67	26	34	0.38	0.34	
1977	6300	241	99	8.98	10	-0.18	4.07	24.5	0.84	0.72	-0.22	-0.40	29	34	0.44	0.40	
1975	7334	182	88	25.35	11	-12.73	-1.30	27.5	1.02	0.83	-4.10	6.30	27	33	0.02	0.99	
1981	7646	178	87	16.94	11	-5.46	1.49	5.0	0.19	0.15	7.44	-3.67	26	33	1.00	0.02	**
1974	7958	132	74	13.17	8	-7.68	-2.40	9.4	0.32	0.28	-0.63	1.73	29	34	0.35	0.85	
1982	7958	158	81	19.22	10	-12.17	-2.84	69.9	2.33	1.94	-3.47	4.66	30	36	0.03	0.99	
				366.15	265												

** indicates serious difficulty in obtaining parameter estimates

9.4.5.3 $\{\tau_i\}$ varies exponentially

Model M6:

$$\tau_i = \exp(\theta + \phi i), \qquad i = 1...60.$$

$$\eta_i = \eta.$$

The transition standard deviation varies as the exponential of a linear function of length and the transition mean is constant. The parameters are transformed to

$$h_L = \theta + \phi l_L.$$

$$h_U = \theta + \phi l_U.$$

Parameter estimates are given in Table 9.10.

The overall fit of the model is not good ($\sum\hat{X}^2$ = 522.24 on 265 d.f.). The transition variance of small trout is always greater than that of large trout. The survival probability of small trout is greater than that of large trout in every year except for 1969 and 1983.

Model M7:

A possible disadvantage of model M6 is that it allows many of the small trout to grow much bigger relative to their original length than the fittest large trout. To overcome this, the transition mean can be allowed to vary linearly with length. However, this involves the estimation of six parameters and the problems of obtaining

TABLE 9.10

Parameter estimates for model M6

Year	E	R_1	R_2	χ^2	df	η	h_L	h_U	τ_L	τ_U	g_L	g_U	l_L	l_U	s_L	s_U
1980	722	286	61	33.59	24	-3.31	2.50	0.62	12.18	1.86	0.36	-2.84	24	38	0.59	0.06
1979	1034	334	73	34.64	25	4.76	2.23	1.23	9.30	3.42	0.44	-3.18	24	38	0.61	0.04
1969	1756	403	74	20.43	16	-7.34	1.73	1.47	5.64	4.35	-2.01	-0.93	26	37	0.12	0.28
1978	2068	429	105	45.89	23	-1.41	2.14	1.38	8.50	3.97	-0.07	-2.19	27	39	0.48	0.10
1968	2272	409	92	27.28	17	-3.13	1.60	0.61	4.95	1.84	-0.36	-2.06	26	37	0.41	0.11
1983	2478	425	43	11.77	11	-8.42	1.64	-1.18	5.16	0.31	-2.90	-1.34	25	35	0.05	0.21
1970	2584	460	115	24.49	23	-2.12	2.47	0.91	11.82	2.48	0.14	-2.23	26	38	0.54	0.10
1972	2888	439	112	41.51	14	-0.59	1.74	1.31	5.70	3.71	-0.14	-1.61	24	33	0.47	0.17
1973	3610	412	119	38.55	20	-3.74	2.00	1.38	7.39	3.97	-0.24	-1.27	25	34	0.44	0.22
1976	4234	354	91	46.72	15	-4.34	1.98	1.31	7.24	3.71	0.09	-1.99	26	34	0.52	0.12
1971	4644	308	110	20.71	13	-2.84	1.57	1.30	4.81	3.67	0.06	-1.12	27	33	0.51	0.25
1967	4652	346	116	58.55	14	-2.96	1.92	1.03	6.82	2.80	0.47	-1.40	26	34	0.62	0.20
1977	6300	241	99	19.13	10	2.76	1.51	0.81	4.53	2.25	0.56	-1.08	29	34	0.64	0.25
1975	7334	182	88	31.89	11	-2.74	1.56	0.98	4.76	2.66	0.60	-0.56	27	33	0.65	0.36
1981	7646	178	87	32.89	11	-3.07	1.67	0.87	5.31	2.39	2.10	-1.29	26	33	0.89	0.22
1974	7958	132	74	11.85	8	-3.57	1.63	-0.17	5.10	0.84	0.95	-0.55	29	34	0.72	0.37
1982	7958	158	81	22.35	10	-3.78	1.75	0.74	5.75	2.10	0.63	-0.25	30	36	0.65	0.44
				522.24	265											

convergence and finding the global minimum become considerable. Instead, the transition mean is allowed to vary in a manner regulated by the transition variance, so that only five parameters have to be estimated:

$$\tau_i = \exp(\theta + \phi i), \qquad i = 1...60.$$

$$\eta_i + 1.645\tau_i = \text{constant},$$

The constraint

$$\eta_i + 1.645\tau_i = \text{constant}$$

imposes an approximate upper limit on the length of a trout at t_2, relative to its original length. If the transition variance decreases as the initial length increases, then the transition mean increases as well. This increase in the mean is sufficient to ensure that only a negligible proportion of the small trout grow much larger, relative to their original length, than the "fittest" large trout. Parameter estimates are given in Table 9.11.

The addition of the variable mean improves the overall fit of the model considerably ($\Sigma\hat{\chi}^2 = 422.63$ on 265 d.f.). However, the fit is not as good as that of model M3. The transition variance of small trout is always greater than that of large trout. The transition variance of small trout is usually biologically reasonable, but in many years the transition variance of large trout is too small. In general, the survival probability of small trout is greater than that of large trout, although the addition of the variable mean often reduces the magnitude of the differences between the probabilities. Very few problems are found in estimating the parameters.

TABLE 9.11

Parameter estimates for model M7

Year	E	R_1	R_2	$\hat{\chi}^2$	df	η_{30}	h_L	h_U	η_L	η_U	τ_L	τ_U	g_L	g_U	l_L	l_U	s_L	s_U	
1980	722	286	61	27.38	24	1.12	1.65	-2.00	-5.71	2.70	5.20	0.14	-0.12	-2.23	24	38	0.47	0.10	**
1979	1034	334	73	28.95	25	3.52	1.78	0.73	-0.02	6.33	5.94	2.07	-0.28	-2.03	24	38	0.43	0.12	
1969	1756	403	74	20.42	16	-8.44	1.66	1.46	-9.06	-7.47	5.27	4.30	-2.38	-0.74	26	37	0.08	0.32	
1978	2068	429	105	30.42	23	-3.70	1.69	0.02	-6.73	0.49	5.41	1.02	-0.56	-1.52	27	39	0.36	0.18	
1968	2272	409	92	23.33	17	-4.50	1.11	0.31	-5.77	-3.00	3.04	1.36	-0.82	-1.46	26	37	0.31	0.19	
1983	2478	425	43	12.15	11	-8.95	-0.16	-1.74	-9.72	-8.61	0.85	0.17	-3.13	-1.25	25	35	0.04	0.23	
1970	2584	460	115	14.13	23	-4.66	1.94	-0.13	-10.39	-0.36	6.97	0.88	-0.56	-1.48	26	38	0.36	0.19	
1972	2888	439	112	32.86	14	-0.17	1.23	-0.01	-3.34	0.67	3.43	0.99	-0.34	-1.48	24	33	0.42	0.19	
1973	3610	412	119	38.02	20	-5.42	1.87	1.35	-8.10	-3.77	6.50	3.87	-1.16	-0.43	25	34	0.24	0.39	
1976	4234	354	91	36.80	15	-3.24	1.43	-0.45	-7.46	-1.60	4.20	0.64	-0.10	-1.84	26	34	0.47	0.14	
1971	4644	308	110	19.51	13	-2.63	1.41	0.72	-4.59	-1.24	4.09	2.05	-0.06	-0.99	27	33	0.49	0.27	
1967	4652	346	116	33.37	14	-3.85	1.79	-0.55	-10.66	-1.74	5.99	0.58	-0.26	-0.82	26	34	0.43	0.31	
1977	6300	241	99	14.20	10	1.61	1.08	-0.07	0.61	3.94	2.95	0.93	0.15	-0.72	29	34	0.54	0.33	
1975	7334	182	88	29.67	11	-1.82	1.24	-0.05	-4.54	-0.40	3.46	0.95	0.76	-0.79	27	33	0.68	0.31	
1981	7646	178	87	26.17	11	-1.52	1.31	-0.50	-5.44	-0.35	3.67	0.61	3.26	-2.10	26	33	0.96	0.11	
1974	7958	132	74	10.20	8	-5.99	1.28	-1.07	-8.21	-2.86	3.59	0.34	-0.44	1.51	29	34	0.39	0.82	
1982	7958	158	81	19.14	10	-9.29	1.69	0.55	-9.29	-3.21	5.42	1.73	-1.02	1.55	30	36	0.27	0.82	
				416.72	265														

** indicates serious difficulty in obtaining parameter estimates

9.4.6 $\{s_i\}$ Quadratic

Model M8:

Elliott (pers. comm.) conducted a series of laboratory experiments in which the survival rate of small and large trout is found to be less than that of trout of medium length. Survival probabilities of this form can be incorporated by the parameterisation

$$s_i = \exp(\theta + \phi i + \psi i^2)/\{1 + \exp(\theta + \phi i + \psi i^2)\}, \qquad i = 1...60.$$

The three survival parameters are transformed to

$$g_L = \theta + \phi l_L + \psi l_L^2,$$

$$g_U = \theta + \phi l_L + \psi l_U^2,$$

$$g_{31} = \theta + 31\phi + 31^2\psi.$$

Parameter estimates assuming "quadratic" survival and normal transitions with constant mean and variance are given in Table 9.12.

The overall fit of the model is not good ($\Sigma\hat{\chi}^2$ = 507.92 on 265 d.f.) and in most years, the minimisation routine is stopped when the magnitude of the estimate of one of the survival parameters becomes too large. In nearly every year, the survival probabilities of small and large trout are greater than those of trout of medium length; these results differ greatly from those observed by Elliott. Such parameter estimates arise because they allow the small trout at t_1 to account for all the small trout and a proportion of the medium length trout at t_2 and allow the large trout

TABLE 9.12

Parameter estimates for model M8

Year	E	R_1	R_2	$\hat{X}^2$	df	η	τ	g_L	g_{31}	g_U	l_L	l_U	s_L	s_{31}	s_U	
1980	722	286	61	36.09	24	-5.10	6.33	-1.15	-3.00	1.38	24	38	0.24	0.05	0.80	**
1979	1034	334	73	31.77	25	0.82	5.49	-0.22	-3.00	0.02	24	38	0.44	0.05	0.51	**
1969	1756	403	74	20.50	16	-7.97	4.56	-2.23	-1.89	-0.58	26	37	0.10	0.13	0.36	
1978	2068	429	105	42.02	23	-3.53	4.97	-0.21	-3.00	0.58	27	39	0.45	0.05	0.64	**
1968	2272	409	92	27.80	17	-4.63	2.87	-0.93	-1.70	-0.60	26	37	0.28	0.16	0.35	
1983	2478	425	43	16.53	11	-8.05	0.86	-2.78	-1.97	-1.35	25	35	0.06	0.12	0.21	
1970	2584	460	115	26.31	23	-6.40	5.70	-1.01	-10.00	2.35	26	38	0.27	0.00	0.91	**
1972	2888	439	112	28.73	14	-1.38	2.92	0.59	-2.38	-0.71	24	33	0.64	0.08	0.33	
1973	3610	412	119	37.95	20	-5.37	4.28	0.23	-3.45	1.42	25	34	0.56	0.03	0.80	
1976	4234	354	91	44.96	15	-6.29	4.55	-0.75	-4.00	2.39	26	34	0.32	0.02	0.92	**
1971	4644	308	110	16.34	13	-3.57	2.94	3.44	-4.00	3.40	27	33	0.97	0.02	0.97	**
1967	4652	346	116	46.98	14	-4.08	2.91	10.00	-7.15	1.09	26	34	1.00	0.00	0.75	**
1977	6300	241	99	19.21	10	1.70	2.84	0.55	-3.00	1.94	29	34	0.63	0.05	0.87	**
1975	7334	182	88	32.44	11	-3.61	3.53	0.02	0.03	4.00	27	33	0.51	0.51	0.98	**
1981	7646	178	87	34.25	11	-4.28	3.47	2.88	-0.77	3.50	26	33	0.95	0.32	0.97	**
1974	7958	132	74	21.05	8	-4.37	3.63	-0.51	-0.24	4.00	29	34	0.37	0.44	0.98	**
1982	7958	158	81	24.99	10	-4.38	3.42	0.71	-0.57	3.50	30	36	0.67	0.36	0.97	**
				507.92	265											

** indicates serious difficulty in obtaining parameter estimates

at t_1 to account for the rest of the medium length trout and all the large trout at t_2. If the survival probabilities of small and large trout are low and the survival probability of medium length trout is high, then a very large transition variance is needed to account for the tails of the length distribution at t_2.

9.5 SIMULATION STUDY

This Section describes a set of simulations to investigate the behaviour of some of the length class transition models. In each simulation, a May/June length class distribution, a model parameterisation and a set of "true" parameter values are selected. These are used to generate 1,000 August/September length class distributions, from which 1000 sets of parameter estimates are obtained for the same model parameterisation. Each simulation investigates how well a model estimates its parameters given that it is the true model.

9.5.1 Model M1: Normal Distribution - Constant Mean and Variance

Histograms and correlations of the simulated parameter estimates for model M1, the year 1971, with the true parameter values

$$\eta = -3.96, \quad \tau = 4.40, \quad g_L = -0.91, \quad g_U = -0.11,$$

are given in Figure 9.8 and Table 9.13 respectively. The true parameter values are those estimated from the original 1971 August/September length class data using model M1. The estimates of η, g_L and g_U are highly correlated and their distributions are bimodal. The bimodality occurs because the model and the estimation procedure are very sensitive to the form of the tails of the August/September length class

Fig. 9.8a Simulated estimates of η.

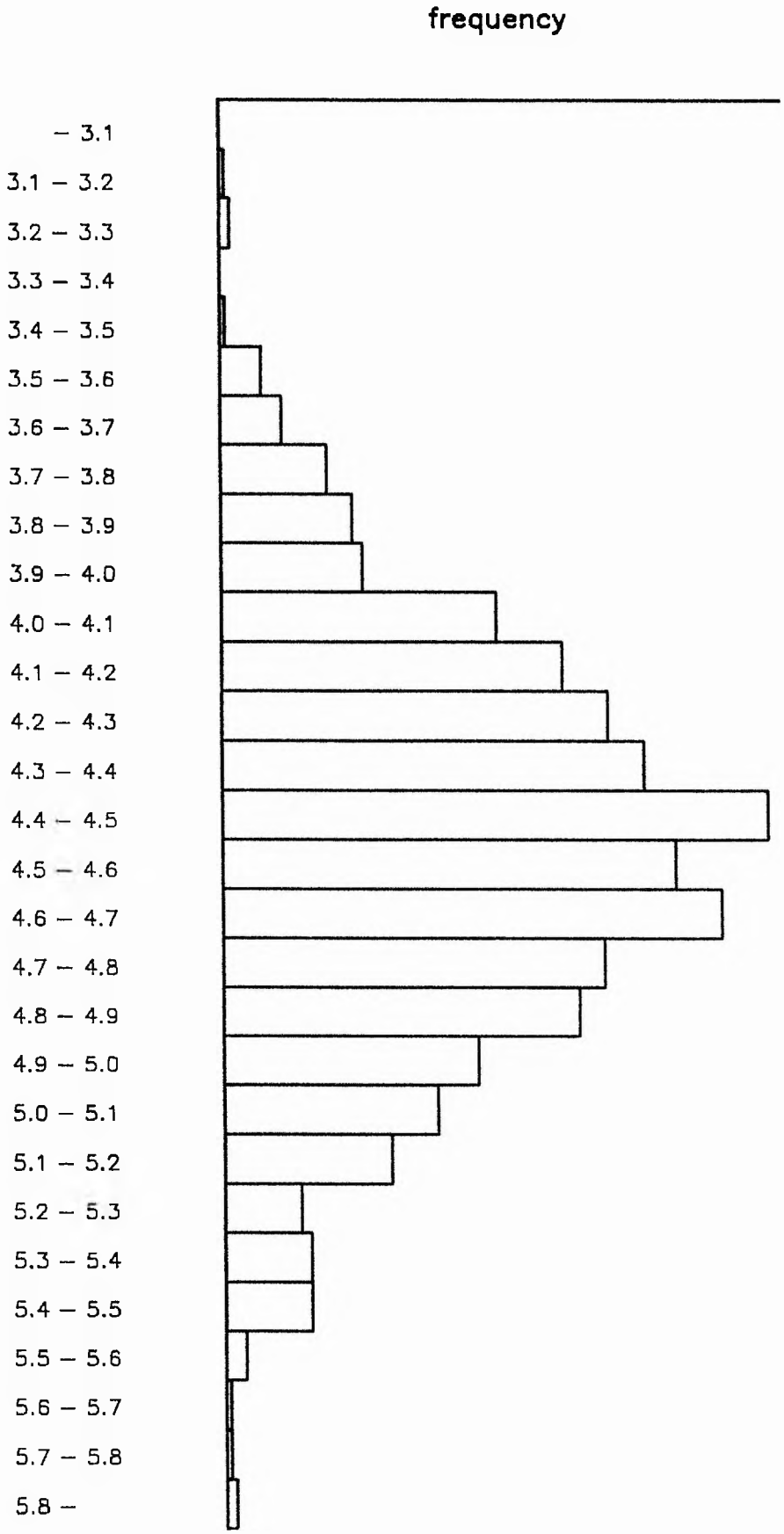


Fig. 9.8b Simulated estimates of τ.

Fig. 9.8c Simulated estimates of g_L.

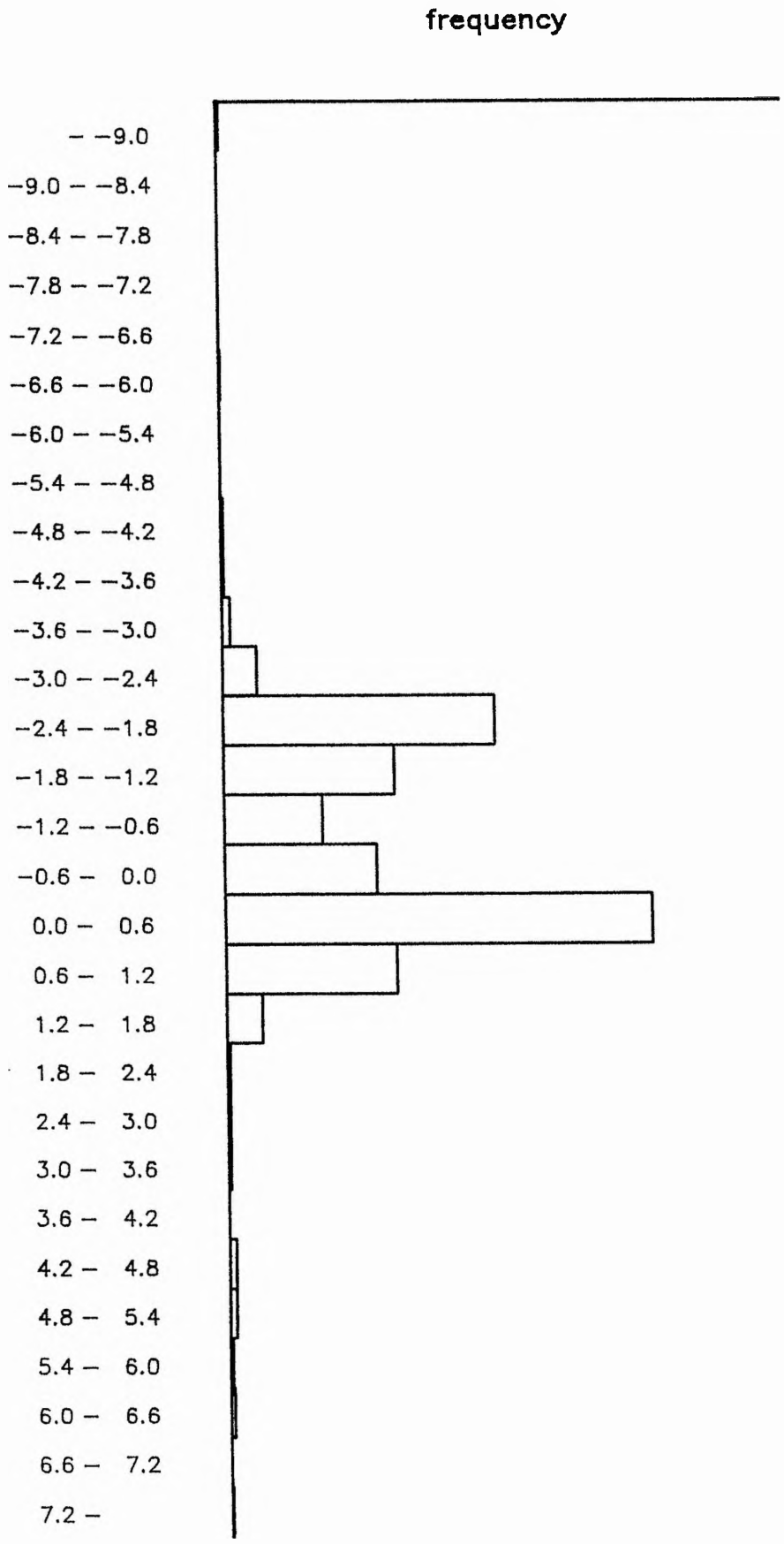


Fig. 9.8d Simulated estimates of g_u.

TABLE 9.13

Correlations between simulated parameter estimates

True values: $\eta = -3.96$, $\tau = 4.40$, $g_L = -0.91$, $g_U = -0.11$

a) η estimated from data

η	1.000			
τ	-0.057	1.000		
g_L	0.795	-0.088	1.000	
g_U	-0.862	0.056	-0.944	1.000

b) η assumed known

τ	1.000		
g_L	-0.100	1.000	
g_U	0.131	-0.847	1.000

distribution. The tails are often quite sparse, and minimising a goodness of fit measure tends to produce one of two types of solution, which give a good fit to one of the tails at the expense of a relatively poor fit to the other. The types of solution are:

i) $\hat{\eta}$ is low, $\hat{g}_L$ is small and $\hat{g}_U$ is large; the survival probability of large trout at t_1 is larger than that of small trout at t_1; the large trout at t_1 account for most of the trout at t_2 so that the transition mean is low (Fig. 9.9a),

ii) $\hat{\eta}$ is high, $\hat{g}_L$ is large and $\hat{g}_U$ is small; the survival probability of large trout at t_1 is smaller than that of small trout at t_1; the small trout at t_1 account for most of the trout at t_2 so that the transition mean is high (Fig. 9.9b).

It is rare that the estimates of η, g_L and g_U take intermediate values since solutions of this form do not give a good representation of either tail.

The bimodality is not a year dependent effect, since three further simulations using the years 1970, 1974, 1978 give similar results. (These years are chosen to cover the full range of initial egg densities). Bimodal distributions are also obtained if the information measure is used instead of the chi-squared measure.

As the true transition variance decreases, the distribution of the parameter estimates becomes unimodal. For example, Figure 9.10 shows a histogram of the simulated estimates of g_L for the year 1971 with $\tau = 2.50$, ($\eta = -3.96$, $g_L = -0.91$, $g_U = -0.11$). (In Fig. 9.10, the estimates of g_L below -9.6 correspond to data sets for which the minimisation routine failed to converge). A small transition variance causes peaks in the May/June length distribution to "transfer" to peaks in the August/September length distribution; the estimation procedure can pair up these peaks and good

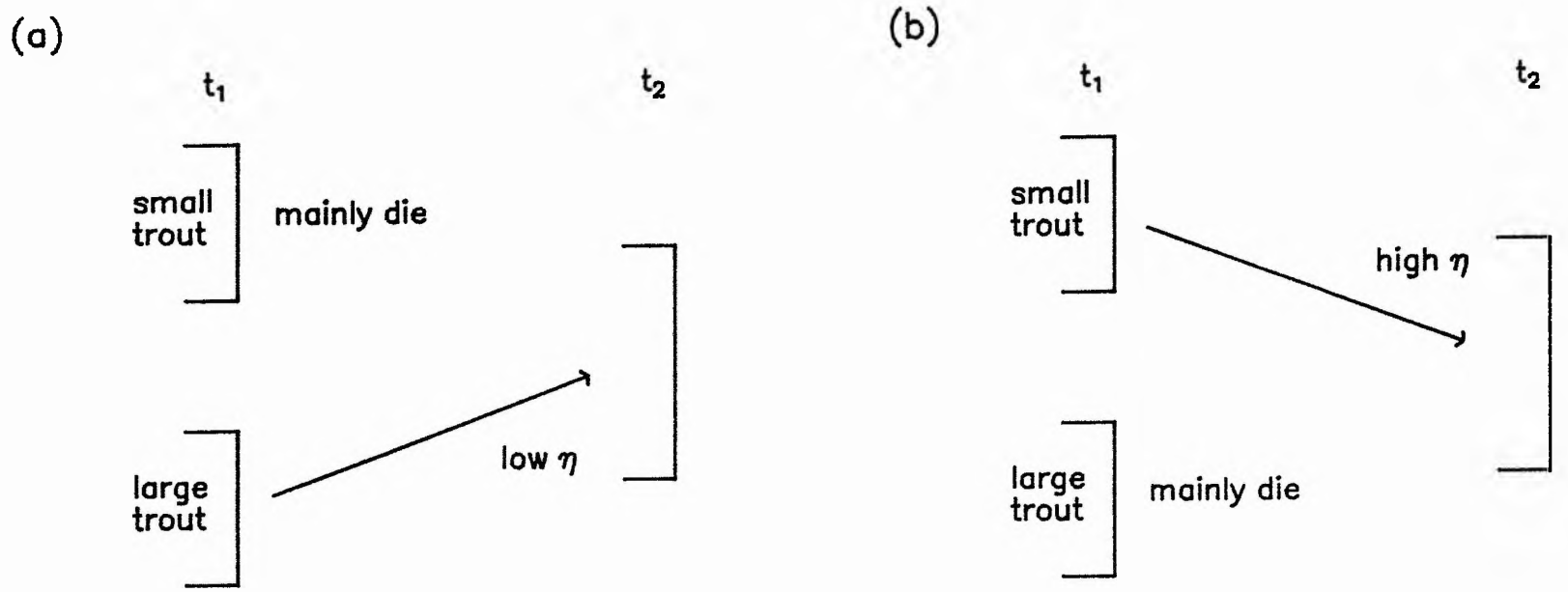


Fig. 9.9 Two types of survival and transition probabilities.

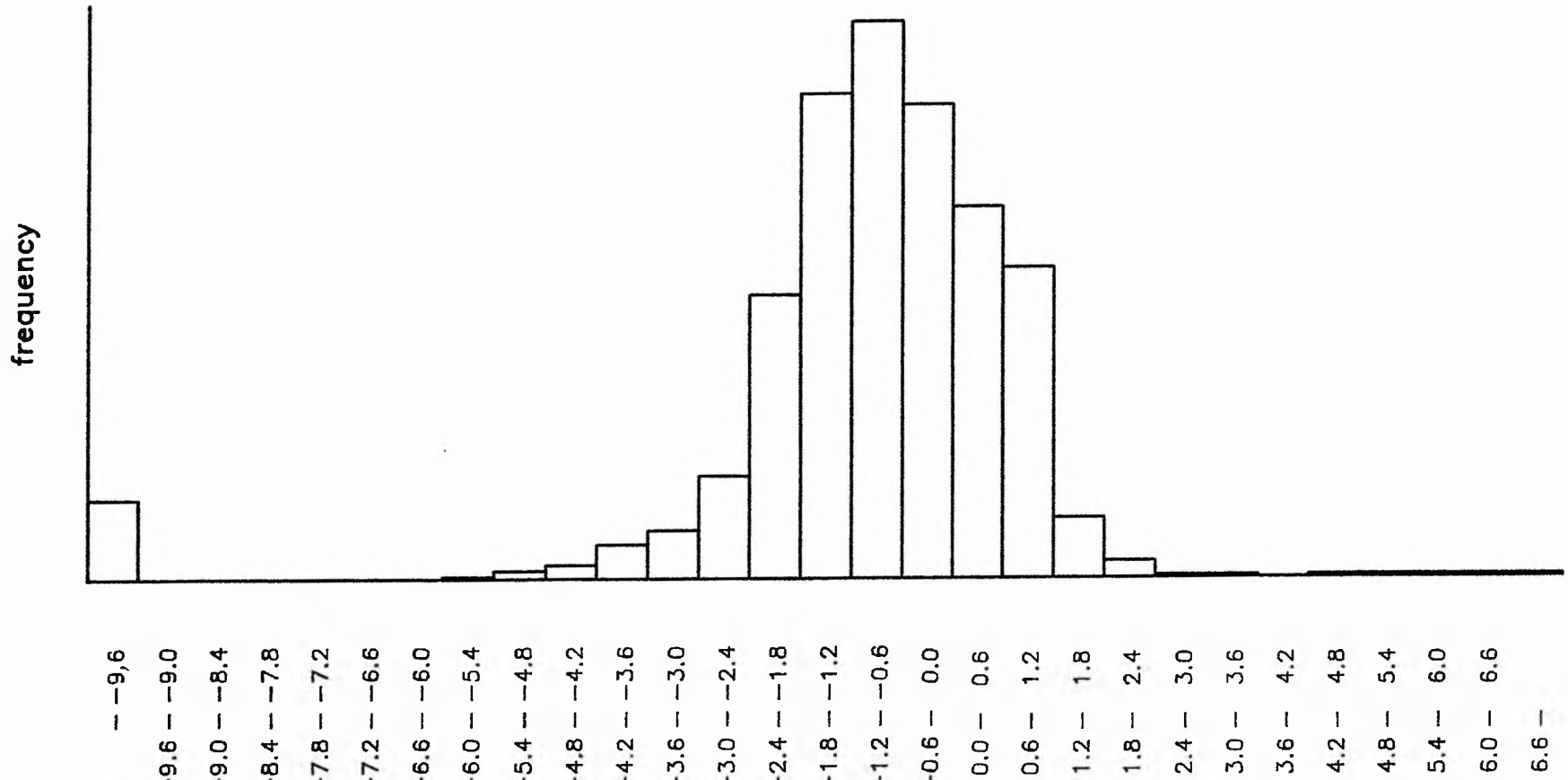


Fig. 9.10 Simulated estimates of g_L when the transition variance is small.

estimates of the survival probabilities are obtained. A large transition variance smooths the peaks in the May/June length distribution and parameter estimation becomes much more difficult.

The extent of the bimodality also decreases if the difference between the survival probabilities of small and large trout is increased. However, numerical difficulties in estimating the parameters greatly increase as well.

Finally, a simulation of the year 1971 with the true parameter values

$$\eta = -3.96, \quad \tau = 4.40, \quad g_L = -0.91, \quad g_U = -0.11,$$

but with η assumed known, shows that the distribution of the parameter estimates is unimodal. For example, a histogram of the simulated estimates of g_L is shown in Figure 9.11. The correlation between the parameter estimates is also reduced (Table 9.13) and although the parameter confidence intervals are large, this can be attributed to the effect of the relatively large transition variance.

9.5.2 Five Parameter Models

Simulations of models M3, M5 and M7 give results that are qualitatively very similar to those obtained above. Provided the transition variances are small, the distributions of the parameter estimates are unimodal and the confidence intervals of the parameters are small.

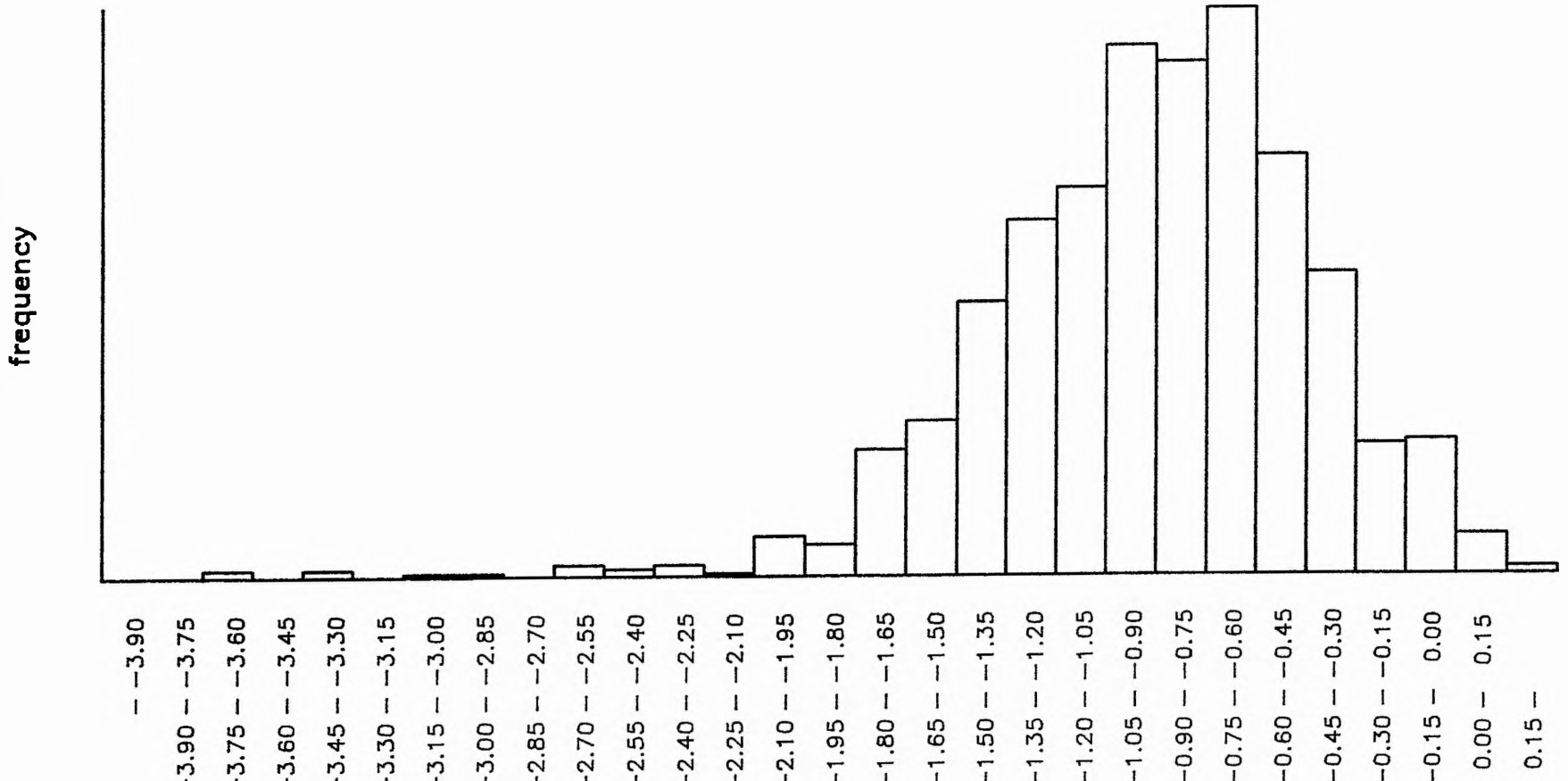


Fig. 9.11 Simulated estimates of g_L when the transition mean is known.

9.6 DISCUSSION

It is not possible to draw any major conclusions from the analysis for two main reasons.

i) Model selection can not be made solely on the basis of the sum of the minimised chi-squared measures. Models which perform well in this respect (models M3 and M5) have the disadvantage that the parameters are hard to estimate and that the estimates are often unrealistic.

ii) The estimates of the survival probabilities are not robust to changes in the parameterisation of the transition probabilities; model M1 indicates that the survival probabilities of large trout are greater than those of small trout, model M7 gives the opposite result and model M3 shows there is no pattern in the survival probabilities. Thus not even qualitative statements can be made about the form of the survival probabilities.

Prior information is needed about the transition or the survival probabilities for the length class transition model to be of use. For example, prior knowledge about the transition means would help in both the selection of an appropriate transition parameterisation and in reducing the problems of parameter estimation if the transition variances are large. Of course, information about transition means is connected with information about transition variances and, implicitly, survival probabilities. Such information would dramatically change the nature of any kind of length class transition model.

CHAPTER 10 POPULATION REGULATION IN BLACK BROWS BECK

The study of population regulation in Black Brows Beck is important for two main reasons. First, there is clear evidence of a relationship between the number of recruits at various stages of the life cycle and the number of eggs E (cf Figs 4.1 and 6.2). Although there have been many studies of other fish populations, few, if any, have provided evidence that stock-recruit relationships exist (Sissenwine, Fogarty and Overholtz, 1988). In many marine species, this is partly because recruitment is likely to be only loosely linked to stock size, except at very low stock values. However, evidence for stock-recruit relationships is also obscured by, for example,

i) large errors in estimating population numbers,

ii) temporal and spatial variation in environmental effects and predation,

(Chapter 2). Such problems can be effectively ignored in the Black Brows Beck study because the study section is comparatively small. For example, environmental factors can be adequately monitored, since it is only necessary to consider a small section of stream. Further, the study section can be extensively electro-fished, so that errors in the estimates of the number or recruits are minimal (Elliott, 1984a). Although there are larger errors in the estimates of the number of eggs, these errors are small compared to the overall range of egg values (Chapter 5) and have a negligible effect on the identification of appropriate stock-recruit relationships. Thus, population regulation in Black Brows Beck can be investigated without having

to consider the "nuisance" factors which are inherent to most other fish population studies.

The second reason is that estimates of population numbers are obtained at regular intervals throughout the life cycle. This allows the identification of

i) stages in the life cycle which are particularly important in regulating population numbers; for example, survival in Black Brows Beck is shown to be density dependent during the first few months of the life cycle (ie to t_2) and density independent thereafter (Chapter 6),

ii) environmental factors which have an important effect on survival; for example, periods of summer drought.

A further advantage is that population models can be developed one stage at a time (cf Chapter 6); in particular, models of the early life stages can be useful in the development of models of the later life stages. For example, the number of 1+ parr at t_4 (ie R_4) can be affected by periods of summer drought in both the first and second summers of the life cycle. The relationship between R_4 and E is well described by model (6.18). However, the form of (6.18) is suggested by the forms of models (6.4) and (6.11) which describe the relationships between R_2 and E and between R_3 and E respectively and in particular, by the relationship between the water level ranges and the corresponding survival rates to t_2 (Section 6.2.5). Had the relationship between R_4 and E been investigated in isolation, it is unlikely that a suitable model such as (6.18) would have been developed.

The relationships between R_1, R_2, R_3, R_4, E^* and E are well described by models (6.6), (6.4), (6.11), (6.18) and (6.21) (Chapter 6). A series of life tables are presented in Table 10. In all but the most serious droughts, initial density is the most important factor regulating the number of fish at any stage of the life cycle. The egg production of a year class E^* is of particular interest in terms of managing a fish stock. In addition to the need to conserve the stock, egg production is important because it is highly correlated with the number of adult fish that return to spawn, and hence with the number of adult fish that can be caught. The form of the stock-recruit relationship (Table 10) shows that stocking the stream would be beneficial only when there is a low number of eggs, ie less than 3,600 eggs in the study section or 60 eggs m^{-2}. Stocking when the number of eggs is greater than this would be counterproductive, since it would reduce the number of adult recruits (Elliott, 1985d). Further, the egg density which maximises the number of adult fish does not depend on the severity of summer droughts which occur whilst the fish are in fresh water. (Presumably, survival has become density independent by the time summer droughts are likely to occur and there is a low but proportional survival rate during a drought with the constant of proportionality depending on the severity of the drought). Thus, even if a drought could be predicted, there is no way that the effect of the drought could be reduced by manipulating the number of eggs in the stream at t_0.

The models of population regulation developed in Chapter 6 are good descriptions of the observed data. However, some of the statistical properties of these models are very complicated, because of the biological constraints that are imposed (eg so that the number of recruits is non-negative). For example, standard linear approximation theory can not be applied to models (6.11), (6.18) and (6.21) (which

TABLE 10

Life Tables

No summer drought (w = 2.97, w' = 2.97)

Eggs	1000	2000	3000	4000	5000	6000	7000	8000	Approx 95% conf limits
R_1	318	424	423	376	313	250	194	148	± 40
R_2	67	101	115	117	111	101	90	78	± 9
R_3	27	41	47	48	45	41	37	32	± 15
R_4	24	37	42	43	41	37	33	29	± 13
E*	3787	5754	6555	6639	6304	5746	5092	4420	± 2911

Moderate drought in first summer (w = 2.28, w' = 2.97)

Eggs	1000	2000	3000	4000	5000	6000	7000	8000	Approx 95% conf limits
R_1	318	424	423	376	313	250	194	148	± 40
R_2	62	94	107	108	103	94	83	72	± 9
R_3	25	38	44	44	42	38	34	29	± 15
R_4	23	35	39	40	38	34	31	27	± 13
E*	3516	5341	6085	6163	5852	5334	4726	4103	± 2911

Serious drought in first summer (w = 1.64, w' = 2.97)

Eggs	1000	2000	3000	4000	5000	6000	7000	8000	Approx 95% conf limits
R_1	318	424	423	376	313	250	194	148	± 40
R_2	25	37	43	43	41	37	33	29	± 9
R_3	10	15	17	18	17	15	14	12	± 15
R_4	9	14	16	16	15	14	12	11	± 13
E*	1403	2131	2428	2459	2334	2128	1885	1637	± 2911

Moderate drought in second summer (w = 2.97, w' = 2.28)

Eggs	1000	2000	3000	4000	5000	6000	7000	8000	Approx 95% conf limits
R_1	318	424	423	376	313	250	194	148	± 40
R_2	67	101	115	117	111	101	90	78	± 9
R_3	27	41	47	48	45	41	37	32	± 15
R_4	20	31	35	35	33	30	27	23	± 13
E*	3108	4721	5379	5448	5172	4714	4178	3627	± 2911

Serious drought in second summer (w = 2.97, w' = 1.64)

Eggs	1000	2000	3000	4000	5000	6000	7000	8000	Approx 95% conf limits
R_1	318	424	423	376	313	250	194	148	± 40
R_2	67	101	115	117	111	101	90	78	± 9
R_3	27	41	47	48	45	41	37	32	± 15
R_4	5	7	8	8	8	7	6	5	± 13
E^*	711	1081	1231	1247	1184	1079	956	830	± 2911

Egg production E^*

Eggs	1000	2000	3000	4000	5000	6000	7000	8000
w = 2.97, w' = 2.97	3787	5754	6555	6639	6304	5746	5092	4420
w = 2.28, w' = 2.97	3516	5341	6085	6163	5852	5334	4726	4103
w = 2.97, w' = 2.28	3108	4721	5379	5448	5172	4714	4178	3627
w = 2.28, w' = 2.28	2885	4382	4993	5057	4801	4376	3878	3367
w = 1.64, w' = 2.97	1403	2131	2428	2459	2334	2128	1885	1637
w = 1.64, w' = 2.28	1151	1748	1992	2017	1915	1746	1547	1343
w = 2.97, w' = 1.64	711	1081	1231	1247	1184	1079	956	830
w = 2.28, w' = 1.64	660	1003	1143	1157	1099	1002	888	771
w = 1.64, w' = 1.64	263	400	456	462	438	400	354	307

describe the relationships between R_3, R_4, E^* and E) because of the change points at ($w_i = \xi$) (cf Section 6.3.1). To some degree, this reflects the fact that the models are developed on the basis of only a small number of observations in drought years. The present study has given some insight into the ways in which drought affects survival. However, it is important to continue to monitor the Black Brows Beck population, particularly in years of low rainfall, so that the drought models can be both modified and refined in the light of new information that becomes available.

Density dependent survival occurs between t_0 and t_1 and between t_1 and t_2. After t_2, survival is density independent (Chapter 6) and there is no evidence that survival depends on any environmental factor other than drought (Elliott, 1985a). Clearly, further work must concentrate on investigating the density dependent mechanisms which govern population regulation between t_0 and t_1. This can be divided into two areas of interest. First, it is important to understand how density affects survival in the first few weeks of the life cycle (ie immediately after the trout have left the gravel bed). This involves a detailed study of, for example,

i) the dispersion of trout from the redd,

ii) the length of time before trout learn to feed,

iii) the ways in which young trout establish and defend territories.

Secondly, the study of survival between t_1 and t_2 also requires additional information. Although the length class model of Chapter 9 does not adequately describe survival between t_1 and t_2, this does not necessarily mean that the survival of an individual fish does not depend on its length. Estimates of the survival probabilities varied greatly with the form of the transition probabilities imposed on the model (Section

9.4). However, this could merely be because the relevant parameters are not estimable from the observed data, rather than because the underlying model is "incorrect". To resolve this question, it is necessary to observe individual trout between t_1 and t_2, to obtain improved estimates of the survival probabilities (Section 9.6). It is possible that a study of survival in the first few weeks of the life cycle will also provide insight into survival between t_1 and t_2. Whatever the true mechanisms behind population regulation in Black Brows Beck, it is likely that future study must concentrate on working at the individual trout level, rather than working with total numbers in the population.

REFERENCES

Bates, D.M. and Watts, D.G. (1980). Relative curvature measures of nonlinearity (with discussion). Journal of the Royal Statistical Society, B, 42, 1-25.

Bates, D.M. and Watts, D.G. (1981). Parameter transformations for improved approximate confidence regions in nonlinear least squares. The Annals of Statistics, 9, 1152-1167.

Beverton, R.J.H. and Holt, S.J. (1957). On the dynamics of exploited fish populations. Fishery Investigations, London Series 2, 19, 1-533.

Box, G.E.P. and Lucas, H.L. (1959). Design of experiments in nonlinear situations. Biometrika, 46, 77-90.

Box, M.J. (1971). Bias in nonlinear estimation (with discussion). Journal of the Royal Statistical Society, B, 33, 171-201

Chambers, J.R. (1973). Fitting nonlinear models: numerical techniques. Biometrika, 60, 1-13.

Cohen, J.E., Christensen, S.W. and Goodyear, C.P. (1983). A stochastic age-structured population model of striped bass (*Morone saxatilis*) in the Potomac River. Canadian Journal of Fisheries and Aquatic Sciences, 40, 2170-2183.

Deriso, R.B. (1980). Harvesting strategies and parameter estimation for an age-structured model. Canadian Journal of Fisheries and Aquatic Sciences, 37, 268-282.

Draper, N.R. and Smith, H. (1981). Applied regression analysis, second edition. John Wiley and Sons, New York.

Elliott, J.M. (1975). The growth rate of brown trout (*Salmo trutta* L.) fed on maximum rations. Journal of Animal Ecology, 44, 805-821.

Elliott, J.M. (1984a). Numerical changes and population regulation in young migratory trout *Salmo trutta* in a Lake District Stream, 1966-83. Journal of Animal Ecology, 53, 327-350.

Elliott, J.M. (1984b). Growth, size, biomass and production of young migratory trout *Salmo trutta* in a Lake District stream, 1966-83. Journal of Animal Ecology, 53, 979-994.

Elliott, J.M. (1985a). Population regulation for different life-stages of migratory trout *Salmo trutta* in a Lake District stream, 1966-83. Journal of Animal Ecology, 54, 617-638.

Elliott, J.M. (1985b). The choice of a stock-recruitment model for migratory trout, *Salmo trutta*, in an English Lake District stream. Archiv fur Hydrobiologie, 104, 145-168.

Elliott, J.M. (1985c). Growth, size, biomass and production for different life-stages of migratory trout *Salmo trutta* in a Lake District stream, 1966-83. Journal of Animal Ecology, 54, 985-1001.

Elliott, J.M. (1985d). Population dynamics of migratory trout, *Salmo trutta*, in a Lake District stream, 1966-83, and their implications for fisheries management. Journal of Fish Biology, 27 (Supplement A), 35-43.

Elliott, J.M. (1986). Spatial distribution and behavioural movements of migratory trout *Salmo trutta* in a Lake District stream. Journal of Animal Ecology, 55, 907-922.

Elliott, J.M. (1987). Population regulation in contrasting populations of trout *Salmo trutta* in two Lake District streams. Journal of Animal Ecology, 56, 83-98.

Elliott, J.M. (1988). Growth, size, biomass and production in contrasting populations of trout *Salmo trutta* in two Lake District Streams. Journal of Animal Ecology, 57, 49-60.

Fox, W.W. (1970). An exponential surplus yield model for optimizing exploited fish populations. Transactions of the American Fisheries Society, 99, 80-88.

Getz, W.M. (1980). The ultimate-sustainable-yield problem in nonlinear age-structured populations. Mathematical Biosciences, 48, 279-292.

Getz, W.M. (1984). Production models for nonlinear stochastic age-structured fisheries. Mathematical Biosciences, 69, 11-30.

Getz, W.M. and Swartzman, G.L. (1981). A probability transition matrix model for yield estimation in fisheries with highly variable recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 38, 847-855.

Hamilton, D.C., Watts, D.G. and Bates, D.M. (1982). Accounting for intrinsic nonlinearity in nonlinear regression parameter inference regions. The Annals of Statistics, 10, 386-393.

Healy, M.J.R. (1984). Review of "Nonlinear regression modeling: a unified practical approach" by D.A. Ratkowsky. Journal of the Royal Statistical Society, A, 147, 531-532.

Hinkley, D.V. (1969). Inference about the intersection in two-phase regression. Biometrika, 56, 495-504.

Hinkley, D.V. (1970). Inference about the change-point in a sequence of random variables. Biometrika, 57, 1-17.

Horwood, J.W. (1982). The variance of population and yield from an age-structured stock with application to the North Sea herring. Journal du Conseil. Conseil permanent international pour l'exploration de la mer, 40, 237-244.

Horwood, J.W. (1983). A general linear theory for the variance of the yield from fish stocks. Mathematical Biosciences, 64, 203-225.

Horwood, J.W. and Shepherd, J.G. (1981). The sensitivity of age-structured populations to environmental variability. Mathematical Biosciences, 57, 59-82.

Hougaard, P. (1985). The appropriateness of the asymptotic distribution in a nonlinear regression model in relation to curvature. Journal of the Royal Statistical Society, B, 47, 103-114.

James, B., James, K.L. and Siegmund, D. (1987). Tests for a change-point. Biometrika, 74, 71-83.

Kendall, M.G. and Stuart, A. (1979). The advanced theory of statistics. Volume 2, fourth edition. Charles Griffin and Co Ltd, London, 748p.

Kimura, D.K., Balsiger, J.W. and Ito, D.H. (1984). Generalized stock reduction analysis. Canadian Journal of Fisheries and Aquatic Sciences, 41, 1325-1333.

Larkin, P.A. and Hourston, A.S. (1964). A model for simulation of the population biology of Pacific salmon. Journal of the Fisheries Research Board of Canada, 21, 1245-1265.

Lowry, R.K. and Morton, R. (1983). An asymmetry measure for estimators in nonlinear regression models. Proceedings of the 44th Session of the International Statistical Institute, Madrid, Contributed Papers, 1, 351-354.

Ludwig, D. and Hilborn, R. (1983). Adaptive probing strategies for age-structured fish stocks. Canadian Journal of Fisheries and Aquatic Sciences, 40, 559-569.

Ludwig, D. and Walters, C.J. (1981). Measurement errors and uncertainty in parameter estimates for stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 38, 711-720.

Ludwig, D. and Walters, C.J. (1985). Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences, 42, 1066-1072.

Marchesseault, G.D., Saila, S.B. and Palm, W.J. (1976). Delayed recruitment models and their application to the American lobster (*Homarus americanus*) fishery. Journal of the Fisheries Research Board of Canada, 33, 1779-1787.

May, R.M. and Oster, G.F. (1976). Bifurcations and dynamic complexity in simple ecological models. American Naturalist, 110, 573-599.

Payne, R.W., Lane, P.W., Ainsley, A.E., Bicknell, K.E., Digby, P.G.N., Harding, S.A., Leech, P.K., Simpson, H.R., Todd, A.D., Verrier, P.J., White, R.P., Gower, J.C., Tunnicliffe Wilson, G. and Paterson, L.J. (1987). Genstat 5 reference manual. Clarendon Press, Oxford.

Pella, J.J. and Tomlinson, P.K. (1969). A generalized stock production model. Bulletin of the Inter-American Tropical Tuna Commission, 14, 421-496.

Ratkowsky, D.A. (1983). Nonlinear regression modeling: a unified practical approach. Marcel Dekker, New York.

Reed, W.J. (1980). Optimum age-specific harvesting in a non-linear population model. Biometrics, 36, 579-583.

Reed, W.J. (1983). Recruitment variability and age structure in harvested animal populations. Mathematical Biosciences, 65, 239-268.

Ricker, W.E. (1954). Stock and recruitment. Journal of the Fisheries Research Board of Canada, 11, 559-623.

Ricker, W.E. (1958). Handbook of computations for biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada, 119.

Roff, D.A. (1983). Analysis of catch/effort data: a comparison of three methods. Canadian Journal of Fisheries and Aquatic Sciences, 40, 1496-1506.

Schaefer, M.B. (1954). Some aspects of the dynamics of population important to the management of the commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Commission, 1, 27-56.

Schaefer, M.B. (1957). A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Bulletin of the Inter-American Tropical Tuna Commission, 2, 245-285.

Schnute, J. (1977). Improved estimates from the Schaefer production model: theoretical considerations. Journal of the Fisheries Research Board of Canada, 34, 583-603.

Schnute, J. (1985). A general theory for analysis of catch and effort data. Canadian Journal of Fisheries and Aquatic Sciences, 42, 414-429.

Shepherd, J.G. (1982). A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. Journal du Conseil. Conseil permanent international pour l'exploration de la mer, 40, 67-75.

Simonoff, J.S. and Tsai, C.-L. (1989). The use of guided reformulations when collinearities are present in non-linear regression. Applied Statistics, 38, 115-126.

Sissenwine, M.P., Fogarty, M.J. and Overholtz, W.J. (1988). Some fisheries implications of recruitment variability. In Fish Population Dynamics, 2nd. ed., The Implications for Management. (J.A. Gulland (ed)). John Wiley and Sons.

Uhler, R.S. (1980). Least squares regression estimates of the Schaefer production model: some Monte Carlo simulation results. Canadian Journal of Fisheries and Aquatic Sciences, 37, 1284-1294.

Walters, C.J. (1985). Bias in the estimation of fuctional relationships from time series data. Canadian Journal of Fisheries and Aquatic Sciences, 42, 147-149.

Walters, C.J. and Ludwig, D. (1981). Effects of measurement errors on the assessment of stock-recruitment relationships. Canadian Journal of Fisheries and Aquatic Sciences, 38, 704-710.

Ware, D.M. (1980). Bioenergetics of stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 37, 1012-1024.

Wolter, K.M. and Fuller, W.A. (1982). Estimation of the quadratic errors-in-variables model. Biometrika, 69, 175-82.

Worsley, K.J. (1986). Confidence regions and tests for a change-point in a sequence of exponential family random variables. Biometrika, 73, 91-104.